

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2002, 13:29:17 / Search time 16 Seconds
(without alignments)
891.882 Million cell updates/sec

Title: US-09-854-562-2
Perfect score: 2519

Sequence: 1 MXSFNTGHNHSTAESGDAY.....SIAGLISSVXTYKPRTHME 485

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/prodata/1/1aa/ECTUS.COMB.pep: *
6: /cgn2_6/prodata/1/1aa/Backlistest.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2519	100.0	485	1	US-08-362-512A-2
2	2519	100.0	485	1	US-08-964-939-2
3	1501	59.6	493	1	US-08-362-512A-4
4	1501	59.6	493	4	US-08-964-939-4
5	567	22.5	447	4	US-09-370-253-6
6	563.5	22.4	447	4	US-09-370-253-10
7	527	20.9	447	4	US-09-370-253-2
8	379	15.0	259	4	US-09-370-253-12
9	310	12.3	432	4	US-09-370-253-12
10	198.5	7.9	473	4	US-09-637-118B-2
11	198.5	7.9	451	1	US-08-191-337-3
12	121.5	4.8	125	4	US-09-370-253-16
13	110	4.4	494	4	US-09-134-001C-4475
14	108	4.3	1212	4	US-09-268-866-2
15	104.5	4.1	438	2	US-08-677-049-9
16	104	4.1	525	4	US-09-351-224E-8
17	102.5	4.1	256	4	US-09-134-001C-6648
18	101.5	4.0	482	4	US-09-134-001C-4309
19	97.5	3.9	2516	3	US-08-374-077C-2
20	97.5	3.9	2516	4	US-08-895-590-2
21	97.5	3.9	2516	4	US-09-539-879A-2
22	97	3.9	467	4	US-09-134-001C-3020
23	97	3.9	489	4	US-09-134-001C-5241
24	97	3.9	521	4	US-09-134-001C-4290
25	96	3.8	518	4	US-09-134-001C-4744
26	95.5	3.8	326	4	US-08-986-768-2
27	95.5	3.8	797	4	US-09-191-468-120

28	95.5	3.8	797	4	US-09-191-468-122	Sequence 122, App
29	93.5	3.7	347	1	US-08-118-270-47	Sequence 47, Appl
30	93.5	3.7	347	5	PCT-US93-08528-47	Sequence 47, Appl
31	93.5	3.7	2522	4	US-09-251-645-13	Sequence 13, Appl
32	93	3.7	424	4	US-09-134-001C-5009	Sequence 5009, Ap
33	92.5	3.7	540	4	US-09-433-994-2	Sequence 2, Appl1
34	92.5	3.7	622	4	US-08-132-990A-4	Sequence 4, Appl1
35	92.5	3.7	622	5	PCT-US92-09382-4	Sequence 4, Appl1
36	92	3.7	109	4	US-09-370-253-8	Sequence 8, Appl1
37	92	3.7	496	4	US-09-134-001C-3001	Sequence 3001, Ap
38	92	3.7	511	2	US-09-073-362-1	Sequence 1, Appl1
39	92	3.7	511	2	US-09-243-920-1	Sequence 1, Appl1
40	90	3.6	349	4	US-09-134-001C-4004	Sequence 4004, Ap
41	90	3.6	797	2	US-08-700-013B-19	Sequence 19, Appl
42	90	3.6	797	2	US-08-700-013B-21	Sequence 21, Appl
43	90	3.6	797	4	US-09-191-468-124	Sequence 124, App
44	89.5	3.6	368	4	US-08-709-838-2	Sequence 2, Appl1
45	89.5	3.6	368	4	US-08-829-839-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-08-362-512A-2
; Sequence 2, Application US/08362512A
; Patent No. 5719043
; GENERAL INFORMATION:
; APPLICANT: FROMMER, Wolf-Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR AN AMINO ACID
; TITLE OF INVENTION: TRANSPORTER, PLASMIDS, BACTERIA, YEASTS AND PLANTS
; TITLE OF INVENTION: CONTAINING A TRANSPORTER AND THEIR USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,512A
; FILING DATE: 05-JAN-1995
; CLASSIFICATION: B03
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/007,636
; FILING DATE: 21-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mellman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/951-107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-512A-2
; Query Match 100.0%; Score 2519; DB 1; Length 485;
; Best local Similarity 100.0%; Pred. No. 1,2e-261;
; Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; 1 MXSFNTGHNHSTAESGDATVSDPTKNVDDEGRKRTGTWLTASAHITTAIVAGSGVSL 60
```

```

;
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-964-939-2

Query Match      100.0%; Score 2519; DB 4; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.2e-261;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 M K S P N T E G N H S T A E S G D A Y T V S D P T K N V D E D G R E K R T G T W L T A S A H I I T A V I G S G V L S L 60
Db 1 M K S P N T E G N H S T A E S G D A Y T V S D P T K N V D E D G R E K R T G T W L T A S A H I I T A V I G S G V L S L 60
Qy 61 A W A I A Q L G W I A G T S I L L I F S F I T Y F T S T M L A D C Y R A P D P V T G K R N Y T Y M D V V R S Y L G R K 120
Db 61 A W A I A Q L G W I A G T S I L L I F S F I T Y F T S T M L A D C Y R A P D P V T G K R N Y T Y M D V V R S Y L G R K 120
Qy 121 V O L C G V A Q Y G N L I G V T V G Y T I T A S I S L V A G K S N C F H D K G H T A D C T I S N Y P Y M A V F G I I Q 180
Db 121 V O L C G V A Q Y G N L I G V T V G Y T I T A S I S L V A G K S N C F H D K G H T A D C T I S N Y P Y M A V F G I I Q 180
Qy 181 V I L S Q I P N F H K L S F L S I M A A V M S F T Y A T I G I G L A I A T V A G G K V G K T S M T G T A V G D V T A A 240
Db 181 V I L S Q I P N F H K L S F L S I M A A V M S F T Y A T I G I G L A I A T V A G G K V G K T S M T G T A V G D V T A A 240
Qy 241 Q K I W R S F Q A V G D I A F A Y A T V L I E I Q D T L R S S P A E N K A M K R A S I L V S L A A A G S I A G L I S S V K T Y K P F 300
Db 241 Q K I W R S F Q A V G D I A F A Y A T V L I E I Q D T L R S S P A E N K A M K R A S I L V S L A A A G S I A G L I S S V K T Y K P F 300
Qy 301 I G Y A A F G N N A P G D F L T D P G F F P F W L I D F A N A C I A V H L I G A Y Q V F A Q P I F Q F V E K K C N R N 360
Db 301 I G Y A A F G N N A P G D F L T D P G F F P F W L I D F A N A C I A V H L I G A Y Q V F A Q P I F Q F V E K K C N R N 360
Qy 361 Y P D N K F I T S E Y S V N V P F L G K F N I S L F R L V W R T A Y V V I T V V A M I P F P F N A I L G I G A A S F 420
Db 361 Y P D N K F I T S E Y S V N V P F L G K F N I S L F R L V W R T A Y V V I T V V A M I P F P F N A I L G I G A A S F 420
Qy 421 W P L T V Y P P V E M H I A Q T K I K K Y S A R W I A L K T M C Y V C L I V S L L A A A G S I A G L I S S V K T Y K P F 480
Db 421 W P L T V Y P P V E M H I A Q T K I K K Y S A R W I A L K T M C Y V C L I V S L L A A A G S I A G L I S S V K T Y K P F 480
Qy 481 R T M H E 485
Db 481 R T M H E 485
```

RESULT 2

```

US-08-964-939-2
; Sequence 2, Application US/08964939
; Patent No. 6245970
; GENERAL INFORMATION:
; APPLICANT: FROMMER, Wolf-Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR AN AMINO ACID
; TRANSPORTER, PLASMIDS, BACTERIA, YEASTS AND PLANTS
; TITLE OF INVENTION: CONTAINING A TRANSPORTER AND THEIR USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,939
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,512
; FILING DATE: 05-JAN-1995
; APPLICATION NUMBER: US 08/007,636
; FILING DATE: 21-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meilman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/951-107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
```

RESULT 3

```

US-08-362-512A-4
; Sequence 4, Application US/08362512A
; Patent No. 5719043
; GENERAL INFORMATION:
; APPLICANT: FROMMER, Wolf-Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR AN AMINO ACID
; TRANSPORTER, PLASMIDS, BACTERIA, YEASTS AND PLANTS
; TITLE OF INVENTION: CONTAINING A TRANSPORTER AND THEIR USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
```

```

:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/362,512A
: FILING DATE: 05-JAN-1995
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/007,636
: FILING DATE: 21-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Mellman, Edward A.
: REGISTRATION NUMBER: 24,735
: REFERENCE/DOCKET NUMBER: P/951-107
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 382-0700
: TELEFAX: (212) 382-0888
: TELEX: 236925
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 493 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-362-512A-4
:
: Query Match 59.6%; Score 1501; DB 1; Length 493;
: Best Local Similarity 56.7%; Pred. No. 2,1e-152;
: Matches 271; Conservative 94; Mismatches 111; Indels 2; Gaps 2;
:
: QY 8 GHNHSTASGDAYVSDPTGNVDEDEGRKRTGTWLTASAHITTAIVIGSGVSLAMALAQ 67
: DB 17 GHQVFDVAHSHFPPQAFKCPDDDLKRTGTWLTASAHITTAIVIGSGVSLAMALAQ 76
: QY 68 GWIAGTSLILFSTFTYFTSTMLADCYRAPDPYTGKRNYYMDVRSYLGGRKYLQCGVA 127
: DB 77 GWIGPAMVLLFSLVLTYSSTLSDCYRTGDVSGKRNYYMDVRSYLGGRKYLQCGVA 136
: QY 128 QYGNLIGVTVYTTASISLVAVGKSNCFHDKGHTADCTISNYPYMAVFGIIOVLSQIP 187
: DB 137 QYLNLFGLAIGYTTAAISMAIKRSNCFHKGNDPCMSNPNMIVFGVAEILLSQVP 196
: QY 188 NFHKLSTLSMAAVMSFTYATIGGLAIVAGKVKTSMTGTAVGVDTAAOKIMSF 247
: DB 197 DFDQIMWSTIAAAMSTYSAIGALIGVAVANGFKSLGISTIGT-VTQOKIMTF 255
: QY 248 QAVGDIAPAVAYATVLEIOTDLRSSPAENKAMKRAISLVGSTFTFFYLGCGLGYAAG 307
: DB 256 QALGDIAFAYSYVLEIOTDLRSSPAENKAMKRAISLVGSTFTFFYLGCGLGYAAG 315
: QY 308 NNAPGDFLTDGFPEFPMILDPANACIAVHLIGAIVFAPQPIFOVEKKCRNRPDNKEI 367
: DB 316 DAAPGNLITGFGFNPFLMDIANAAIVHLVGAIVFAPQPIFAFIEKSVARERPDNDFL 375
: QY 368 TSEYSVNVP-FLGKFNISLRLVWRTAYVYITTVAMIPFPFNALIGLIGAASFPLTVY 426
: DB 376 SKEPEIRIRPGFKSPYKNVFRMYRVSFVTTTIVISMLMPFNDVVGILGALGFPLTVY 435
: QY 427 FPEVMEHIAQTIKKYKSAWIAKTMCVCLIVSLAAGSIAGLISSVKTYPFRTH 484
: DB 436 FPEVMEYIKQKVERKSTRWVCLQMLSVACIVSVAGVSIAGVMDLKVYKPFKSTY 493
:
: RESULT 4
: US-08-964-939-4
: Sequence 4, Application US/08964939
: Patent No. 6245970
: GENERAL INFORMATION:
: APPLICANT: FROMMER, Wolf-Bernd
: TITLE OF INVENTION: DNA SEQUENCES FOR AN AMINO ACID
: TITLE OF INVENTION: TRANSPORTER, PLASMIDS, BACTERIA, YEASTS AND PLANTS
: TITLE OF INVENTION: CONTAINING A TRANSPORTER AND THEIR USE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Oetzielen, Faber, Gerb & Solfen
: STREET: 1180 Avenue of the Americas

```

```

:
: CITY: New York
: STATE: NY
: COUNTRY: US
: ZIP: 10036-8403
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/964,939
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/362,512
: FILING DATE: 05-JAN-1995
: APPLICATION NUMBER: US 08/007,636
: FILING DATE: 21-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Mellman, Edward A.
: REGISTRATION NUMBER: 24,735
: REFERENCE/DOCKET NUMBER: P/951-107
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 382-0700
: TELEFAX: (212) 382-0888
: TELEX: 236925
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 493 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-964-939-4
:
: Query Match 59.6%; Score 1501; DB 4; Length 493;
: Best Local Similarity 56.7%; Pred. No. 2,1e-152;
: Matches 271; Conservative 94; Mismatches 111; Indels 2; Gaps 2;
:
: QY 8 GHNHSTASGDAYVSDPTGNVDEDEGRKRTGTWLTASAHITTAIVIGSGVSLAMALAQ 67
: DB 17 GHQVFDVAHSHFPPQAFKCPDDDLKRTGTWLTASAHITTAIVIGSGVSLAMALAQ 76
: QY 68 GWIAGTSLILFSTFTYFTSTMLADCYRAPDPYTGKRNYYMDVRSYLGGRKYLQCGVA 127
: DB 77 GWIGPAMVLLFSLVLTYSSTLSDCYRTGDVSGKRNYYMDVRSYLGGRKYLQCGVA 136
: QY 128 QYGNLIGVTVYTTASISLVAVGKSNCFHDKGHTADCTISNYPYMAVFGIIOVLSQIP 187
: DB 137 QYLNLFGLAIGYTTAAISMAIKRSNCFHKGNDPCMSNPNMIVFGVAEILLSQVP 196
: QY 188 NFHKLSTLSMAAVMSFTYATIGGLAIVAGKVKTSMTGTAVGVDTAAOKIMSF 247
: DB 197 DFDQIMWSTIAAAMSTYSAIGALIGVAVANGFKSLGISTIGT-VTQOKIMTF 255
: QY 248 QAVGDIAPAVAYATVLEIOTDLRSSPAENKAMKRAISLVGSTFTFFYLGCGLGYAAG 307
: DB 256 QALGDIAFAYSYVLEIOTDLRSSPAENKAMKRAISLVGSTFTFFYLGCGLGYAAG 315
: QY 308 NNAPGDFLTDGFPEFPMILDPANACIAVHLIGAIVFAPQPIFOVEKKCRNRPDNKEI 367
: DB 316 DAAPGNLITGFGFNPFLMDIANAAIVHLVGAIVFAPQPIFAFIEKSVARERPDNDFL 375
: QY 368 TSEYSVNVP-FLGKFNISLRLVWRTAYVYITTVAMIPFPFNALIGLIGAASFPLTVY 426
: DB 376 SKEPEIRIRPGFKSPYKNVFRMYRVSFVTTTIVISMLMPFNDVVGILGALGFPLTVY 435
: QY 427 FPEVMEHIAQTIKKYKSAWIAKTMCVCLIVSLAAGSIAGLISSVKTYPFRTH 484
: DB 436 FPEVMEYIKQKVERKSTRWVCLQMLSVACIVSVAGVSIAGVMDLKVYKPFKSTY 493
:
: RESULT 5
: US-09-370-253-6

```

```
; Sequence 6, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-370-253-6

Query Match      22.5%; Score 567; DB 4; Length 447;
Best Local Similarity 31.9%; Pred. No. 3.2e-52;
Matches 153; Conservative 81; Mismatches 167; Indels 78; Gaps 18;

Qy 31 EDGR-----EKRTGTWLTASAHITAVIGSVLSLAWAIAQLGNIAGTSIL 76
Db 13 KGRSAQAEKAIDWLPTITSSRNKMWYSAFHNVTAMVGAGVLSLPSYAMSELGNGPGIAVL 72
Qy 77 LIFSFTVFTSTMLADCYRAPDPVTGKRNVTYMDVRSYLGGRKVQLCGVA--QVGNLIG 134
Db 73 ILSWIIITLYTLQWVEMH---EMVPGKRPDRYHE-LGQHAFGKLGKLVIVPQQLVVEVG 128
Qy 135 VTGVYTTTASISLAVAGKSNCFHD---KGHTADCTISNYPYMAVFGIIQVILSQIDPNFH 191
Db 129 VNIYVMVTGKSL-----KKFHDVLCGCGCKNIKLTV-FIMIFASVHFLSQLPNFS 181
Qy 192 LSPISMAAVNSFTYATIGTIGLAIATVAGGKVGKTSMTGTAVGVDDVTAQAQKIMRSFOAVG 251
Db 182 ISGVSLSAAVNSLSYSTIANG---ASVDKGMVADVDYHLRA-----TTSTGKVGFFSALG 234
Qy 252 DIAPAYAYATVLEIQTLLRSSPAENKAMKRASLVGVSTTTFFYLCL----GCIGYAAFG 307
Db 235 DVAFAYAGHNVLIEIQATIPSTP--EKPSKPKMKGVVYIIVVVALCFYFVALGYWAFG 292
Qy 308 NNAPGDFLTDGFFPEPFWLIDFANACIAVHLIGAYQVPAQIFQVE-----KKNRNYPD 363
Db 293 NHVDNLTLS--RPNKLIALANMMVVIHVIGSYQIYAMPVDFMDIETVLVKKL--RFP- 347
Qy 364 NKFITSEYVNVVPLGKFNISLRLVWRTAYVITTVVAMIPPPFNAILGLIGAASFWPL 423
Db 348 -----PGL-----TLRLIARTLYVAFTMEVAITFPFFGGLGFFGFGFAFPT 389
Qy 424 TVYFPVEMHIAQTKIKYSARWIALKTMCVCLIVS---LLAAAGSIAGLISSVKTYK 478
Db 390 TYFLPCIMWLAIYKPRFSLSWFT---NWCIVLGVCLMILSPIGGLRQIILDAKTYK 444

RESULT 6
US-09-370-253-10
; Sequence 10, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97

; SEQ ID NO 10
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-370-253-10

Query Match      22.4%; Score 563.5; DB 4; Length 447;
Best Local Similarity 32.1%; Pred. No. 7.5e-52;
Matches 154; Conservative 85; Mismatches 168; Indels 73; Gaps 16;

Qy 25 PTRKVDDEGREK-----RTGTWLTASAHITAVIGSVLSLAWAIAQLGNIAGTS 74
Db 12 PAKDVRTDEQKIDDLWLPITSSRNKMWYSAFHNVTAMVGAGVLSLPSYAMSELGNGFGIA 71
Qy 75 ILLIFSFTVFTSTMLADCYRAPDPVTGKRNVTYMDVRSYLGGRKVQLCGVAQYGNL-I 133
Db 72 VLIVSWITLYTLQWVEMH---EMVPGKRPDRYHELGQHAFGKLGKLVIVPQQLIVEV 128
Qy 134 GTVGYTTTASISLAVAGKSNCFHD---KGHTADCTISNYPYMAVFGIIQVILSQIDPNFH 190
Db 129 GVNIVYVMVTGGRSL-----KKFHDVLCGCGCKDIKLTFF-FIMIFASVHFLSQLPNLN 180
Qy 191 KLSPLSMAAVNSFTYATIGTIGLAIATVAGGKVGKTSMTGTAVGVDD---VTAQAQKIWR 245
Db 181 SISGVSLSAAVNSLSYSTIANG---ASVDKGM-----VVDVYNLRATTTPGKVFG 228
Qy 246 SFQAVGDIAPAYAYATVLEIQTLLRSSPAE--NKAMKRASLVGVSTTTFFYLCCGIGY 303
Db 229 FFGALGEVAPAYAGHNVLIEIQATIPSTPEKPSKPKMKGVVYIIVVVALCFYFVALIGY 288
Qy 304 AAFNNAPGDFLTDGFFPEPFWLIDFANACIAVHLIGAYQVPAQIFQFVEKKKNRNYPD 363
Db 289 WAFNSVDDNLTILN--KPTWLIATANMMVVIHVIGSYQIYAMPVDFMDI----- 337
Qy 364 NKFITSEYVNVVPLGKFNISL-FRLVWRTAYVITTVVAMIPPPFNAILGLIGAASFWP 422
Db 338 -----TVLVKKL-KFPFGLTLRLIARTLYVAFTMEVAITFPFFGGLGFFGFGFAFAP 388
Qy 423 LTVYFPVEMHIAQTKIKYSARWIALKTMCVCLIVS---LLAAAGSIAGLISSVKTYK 478
Db 389 TTYPCLPCIMWLAIYKPRFSLSWFT----NWCIVLGVCLMILSPIGGLRQIILDSKTYK 444

RESULT 7
US-09-370-253-2
; Sequence 2, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Hordeum vulgare
US-09-370-253-2

Query Match      20.9%; Score 527; DB 4; Length 447;
Best Local Similarity 28.9%; Pred. No. 6.2e-48;
Matches 142; Conservative 89; Mismatches 177; Indels 84; Gaps 18;

Qy 21 TVSDPTKKNV-----DEGREKRTGTWLTASA-----HIITAVIGSVLSLAWAIA 65
Db 3 THASPDNPTPKDERTAREKAIDDLPTITSSRNKMWYSAFHNVTAMVGAGVLSLPSYAMS 62
```

```

Qy 66 QLGWLAGTLLIPSFITYETSMADCYRAPDPYKKNYTMVRSYLGRRKVLGC 125
Db 63 ELGNGPGLAVWTLSMTITVTTLWQVEMH---EMVGNKFDREYHEL-----G 106
Qy 126 VAQYGNLIG--VTYGYTTTASISL-----VAQGS-NCPHD---KGTADCTISNYPMA 174
Db 107 QHAFGDKLGMIVVPOOLVVEVSLNIVYWTGNSLKPHDVIQCGKCKDIKLTV--FIM 164
Qy 175 VEGIIQVLSIQIPNPHLSPLSTMAVMSFTATTCIGALITVAGKVGKTSMTGAVG 234
Db 165 IFASVHVLSOLPFPNFSISGISTAAVMSLSYTIMWGSLSL-----HKXEBNDVSLR 218
Qy 235 VDTVAQKIMRSFOAVGDIAFAAYAVATVLEIETDLRSSPA--ENKAMKRAISLVGSTTT 292
Db 219 ASATAGQ-VFQFLQGLGDVAFSGHNVVLEIQATIPSTPGNSKKPMKGVVAYIIIA 277
Qy 293 FFYLLQCGICGAAPGNNAPELFDGFFPEPFLIDFANACIAVHLIGAYQVFAPIFOF 352
Db 278 ACTPVPVAFIGWAGNSVDNIIITLN--KPKMLIMANMMVVLIGSYQIYAPVFD 335
Qy 353 VEKKONRNPDKRFTS--EYSVNVPELGKFNISLFLVWRTAYVVTTVVAMTPEPENA 410
Db 336 ME-----FLVKKLEFARGI-----TLRLITRTIYVAFMTPIGMSPPFGG 376
Qy 411 ILGLIGASFPWPLVYFPVPMHIAOKIKKYSARWIALKTMCYVCLISL-----AAAGS 466
Db 377 LIGFFGLAPFITYFLPCIMWLITCKPRIPLSWFT---NMICVLGVLMLTVAPIG 432
Qy 467 IAGLISSVKT 478
Db 433 LRQIISAKTYK 444

```

RESULT 8
US-09-370-253-4
Sequence 4, Application US/09370253
Patent No. 6165792

```

GENERAL INFORMATION:
APPLICANT: Allen, Stephen M.
APPLICANT: Sakai, Hajime
APPLICANT: Thorpe, Catherine J.
TITLE OF INVENTION: Amino Acid Transporters
FILE REFERENCE: BB-1200
CURRENT APPLICATION NUMBER: US/09/370,253
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: 60/097,222
EARLIER FILING DATE: August 20, 1998
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Microsoft Office 97
SEQ ID NO 4
LENGTH: 259
TYPE: PRT
ORGANISM: Zea mays
US-09-370-253-4

```

Query Match 15.0%; Score 379; DB 4; Length 259;
Best Local Similarity 33.9%; Pred. No. 2, 1e-32;

Matches 99; Conservative 44; Mismatches 101; Indels 48; Gaps 11;

```

Qy 199 AAVMSFTYATIGLAIATVAGKVGKTSMTGAVD-----VTAQKIMRSFOAVDI 253
Db 1 AAVMSLSYSTIAMG--ASVHKGRMS-----GVGYHLRATTTGKVFEGFGALGDV 48
Qy 254 AFAYAVATVLEIETDLRSSP--AENKAMKRAISLVGSTTTFFYLLQCGICGAAPGNNA 311
Db 49 AFAYAGNVVLEIQTIPSTPDKSKPMKGVVAVVVALCYFPVALIGWAGNIVE 108
Qy 312 GDFLTDFGFFPEPFLIDFANACIAVHLIGAYQVFAPIFOVE---KKONRNPDKRFT 367
Db 109 DNIIITLS--KPKMLIALANMMVVLIGSYQIYAMPVDMIEIVLTKL--RFP----- 159
Qy 368 TSESVNVNPELGKFNISLFLVWRTAYVVTTVVAMTPEPFAIIGLIGASFPWLT 427

```

```

Db 160 -----PGL-----TLRLIARTLVAFMTFIAITPPFPGSLGFPFGAFAPATTYFL 205
Qy 428 PEVEMHIAQTKIKKYSARWIALKTMCYVC-LIYSLLAAGSITAGLISSVKT 478
Db 206 PCVWMLAIYKPKRFSLSWL--NMCTIILGVLMLTSPIGRLQIIMDAKTYQ 256

```

RESULT 9
US-09-370-253-12
Sequence 12, Application US/09370253
Patent No. 6165792

```

GENERAL INFORMATION:
APPLICANT: Allen, Stephen M.
APPLICANT: Sakai, Hajime
APPLICANT: Thorpe, Catherine J.
TITLE OF INVENTION: Amino Acid Transporters
FILE REFERENCE: BB-1200
CURRENT APPLICATION NUMBER: US/09/370,253
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: 60/097,222
EARLIER FILING DATE: August 20, 1998
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Microsoft Office 97
SEQ ID NO 12
LENGTH: 432
TYPE: PRT
ORGANISM: Zea mays
US-09-370-253-12

```

Query Match 12.3%; Score 310; DB 4; Length 432;
Best Local Similarity 25.4%; Pred. No. 1, 1e-24;
Matches 123; Conservative 71; Mismatches 216; Indels 74; Gaps 19;

```

Qy 14 AESGDATVSDPTQVNDVEDEREKRTGTWLTASAHITTVAGISGVLSLMAI-AQLGWAG 72
Db 3 AAADKREISDDT-----AQISVDPMYQVGFVLITGNSAYVIGYSGSITWPLGWTG 56
Qy 73 TSILIFSFIYFTSTMLADCYRAPDPYTKKNYTMVRSYLGRRKVLGCYAOYGNL 132
Db 57 TCGILLAAISMTANMLARLH-----VGGKHIRYRDLAGHIYPRKIYGLTMLQYINL 112
Qy 133 IGVTVGYTTTASISLVAVGKSNCFHDKGTADCTISNIPY-MAVGCIIQVLS-QIPNH 190
Db 113 FMINTGFIILAGQALKAT--YGLFSDG-----VLKJPCYLAISGFYCALFAIGIPLS 164
Qy 191 KLSFLSTMAVMSFTYATIGLAIATVAGKVGKTSMTGAVDVAQKIMRSFOAV 250
Db 165 ALRIMLGSTILFSLMTYIVAVLS-----SRDGIAPARQYSIKSSQSTRVFTTI 215
Qy 251 GDIA-FAYAT-VLIEIQTDLRSSPAENKAMKRAISLVGSTTTFFYLLQCG----- 300
Db 216 GSIDLVFAVNTGMLPEIQATIR--PPVKNMEKA-----LMFOPTISLPLYAVVF 265
Qy 301 IGVAFAFNNAPELFDGFFPEPFLIDFANACIAVHLIGAYQVFAPIFOVEKKCN 360
Db 266 VGWMAVSGSTSGYILN--SVTGRVWKAVANLSAPFOYIALHITASMYEFLDK----- 319
Qy 361 YPDNKFITSESVNVPELGKFNIS--LFLVWRTAYVVTTVVAMTPEPFAIIGLIGAA 418
Db 330 YGSGG-----GPEIHNVAFVAVAGGLITVNTLVAAVLEFLDDPMSLTGAL 367
Qy 419 SFWPLTVYFPEVEMH--IAQTKIKKYSARWIALKTMCYVCLISLAAGSITAGLISSVKT 476
Db 368 STFPITFLANHWLAWKPKRLGAIQKSHWNLVIGFTALAA--AAVSALIRLIRDSST 425
Qy 477 YKPF 480
Db 426 YHFF 429

```

RESULT 10
US-09-637-118B-2
Sequence 2, Application US/09637118B

Patent No. 6423838
GENERAL INFORMATION:
APPLICANT: Hitachi LTD
TITLE OF INVENTION: A gene for proline transporter in rice
FILE REFERENCE: NITT.0008,NT0114US
CURRENT APPLICATION NUMBER: US/09/637,118B
CURRENT FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: JP 230291
PRIOR FILING DATE: 1999-08-17
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent in version 3.0
SEQ ID NO 2
LENGTH: 473
TYPE: PRT
ORGANISM: Oryza sativa L.
FEATURE:
NAME/KEY: CDS
LOCATION: 1..473
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: BAA93437
DATABASE ENTRY DATE: 1999-01-25
US-09-637-118B-2

Query Match 12.3%; Score 310; DB 4; Length 473;
Best Local Similarity 24.4%; Pred. No. 1.3e-24;
Matches 123; Conservative 84; Mismatches 205; Indels 92; Gaps 19;

Qy 23 SDPTKNDVDEGKRTGTMLTA-----SAHIITA-----VIGSGVL 58
Db 19 SAPSSLGADGEERETVPLLSCKMADKSDTVQVSEDTAHQISIDPMYQVGFILTTGVN 78
Qy 59 SL-----ANAIAQLGIACTSLILFSPITFTSTMLADCYRPPDVTGKNTYMDVV 112
Db 79 SAYVLGYSASIMPLGWTGCTGLILAAISMYANALLAHLHE---VGGKRHRIRYDLA 134
Qy 113 RSYLGGKRVQLCGVAQGNLIGVTGTYTITAGISLVAVGKSNCFHDKGHTADCTISNYPY 172
Db 135 GHYGRKMSYLTWALQYVNLFMINTGTLILAGALKAI--VYLFEDDG-----VLKLPY 186
Qy 173 -MAVFGIIQVILS-QIPNFKLFSLSMAVMSFTYATIGIGLAIAIVAGKVKTSMTG 230
Db 187 CIALSGFVFCALFAGIPYLSALRIWGLSTVFSLIYIMAFVMSL-----RDGITT 237
Qy 231 TAVGVDV--TAAQKIKRSFOAGDIAFAYAYATVLIETQDILRSSPAEN--KAMKRASLV 286
Db 238 PAKDYTIPIGSHSDRIFTTIGAVANLVFAY-NTGMLPEIQATIRPPVKNMEKALWFOFTV 296
Qy 287 G---VSTTTTPYLICGIGYAAACGNAPGDFLTDGFPPEPFLIDFANACIAVHLIGAYQ 343
Db 297 GSLPLYAVTF-----MGYWAYGSTSYLLNSVK--GPIMIKTVANLSAFLQTVIALH 347
Qy 344 VFAQPIFQFVEKKKNRNPDKNETITSEYVNVFPFGCKFNIS--LPRLVWRVAYVVIITVV 401
Db 348 IFASPMVEFLDTREGSGHG-----GPAIHNIMFRVGVGSLYLVNTLV 391
Qy 402 AMIPFPFNAILGLIGAASFPLTYVPFVEMH--IAQTKIKKYSARWIAKTMVCVCLIVS 459
Db 392 AAMLPFLGDFMSLTGALSTFPLFLVLANHMYLTUVXQNKMSIFRCKMHWLVNVPFSCIL--S 449
Qy 460 LLAAGAGSIAGLISSVKTYKPFRTM 483
Db 450 VAAAVAARLITVDYSTYHLFADM 473

RESULT 11
US-08-191-337-3
; Sequence 3, Application US/08191337
; Patent No. 5643745
; GENERAL INFORMATION:
; APPLICANT: STUART, W. DORSEY
; TITLE OF INVENTION: HETEROLOGOUS DIMERIC PROTEINS PRODUCED
; IN HETEROKARYONS
; NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/191,337
FILING DATE: 03-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 3918-0002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 451 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-191-337-3

Query Match 7.9%; Score 198.5; DB 1; Length 451;
Best Local Similarity 22.0%; Pred. No. 1.1e-12;
Matches 102; Conservative 74; Mismatches 175; Indels 113; Gaps 23;

Qy 31 EDGREK--RTGWTLTASAHITAVIGSGVLSLAIAQLGIACTSLILFSPITFTST 88
Db 44 QEGHAKFRLG-WKRLTVLIVEALALGSLPCGAFATLGMVPGVILSVGMLICIYTAH 102
Qy 89 MLADCYRPPDVTGKNTYMDVRSYLGGRKVLQCGVAQGNLIGVTGTYTITASISLV 148
Db 103 VIGOTKLXHEIA-----HYADLGRVWFRMGVEIISPMFVLQILFIVGSHVLTGIMWG 157
Qy 149 AVGKSNCFHDKGHTADCTISNYPYMAVFGIIQVI-----LSQIPNFHK-----LSFLSIM 198
Db 158 TI-----TUNGN-GTCSL-----VFGIVSAIILPLAIPPSFAEAVAILGYIDFVSIC 203
Qy 199 AAVMSFTYATIGIGLAIAIVAGKVKTSMTGTAVGV-----DVTAAQKIWRSPQAV 250
Db 204 AAIL-----ITMIATG-IRSSHQEGGLAAVPWSCPKEDLSLAE----GPIAV 246
Qy 251 GDIAPAVAVATVLIETQDILRSSPAENKAMKRASLVGVSTTTFFYLICGIGYAAFGNNA 310
Db 247 SNIVFAYSPMCOFSPMDEMHTPSDYKKSIALGLIEI----FYITVGGVYVAFVGPV 302
Qy 311 PGDFLTD-----FGFFBFWLIDFANACIAVHLIGAYQVFAQPIFQFVEKKKNRY 361
Db 303 QSPALLSAGPLAKAVFALP---VIFISGINTVVSRYLI-----ERIW 346
Qy 362 PNKFTITSEYVNVV-----PLGKFNISLFRVLVWRVAYVVIITVVAMIFFPFAAILGLI- 415
Db 347 PNN---VIRY-VNTPAGWVWVLG-FDFGITLIAM-----VIAEAIFFPSDLAICS 392
Qy 416 -----GNAEFWPLTYVPFVEMHIAQTKIKKYSARWIAKTMVCYV 454
Db 393 ALFISGFSFYFPALMYFKITRNDAKSQGKYFLD--ALNMLCFV 434

RESULT 12
US-09-370-253-16
; Sequence 16, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2002, 13:28:42 ; Search time 21 Seconds

(without alignments)
2220.247 Million cell updates/sec

Title: US-09-854-562-2

Perfect score: 1 MMSFTEGHNHSTASGDAY.....SIAGLISSVKTYKPFRTME 485

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2519	100.0	485	2	A48187 amino acid transpo
2	1974.5	78.4	475	2	T00620 probable amino aci
3	1871	74.3	481	2	T50691 amino acid permeas
4	1589.5	63.1	385	2	T07130 probable amino aci
5	1540.5	61.2	486	2	T10100 amino acid transpo
6	1504	59.7	476	2	H96802 probable amino aci
7	1501	59.6	493	2	S82421 amino acid transpo
8	1406.5	55.8	476	2	C96505 probable amino aci
9	1326	52.6	376	2	T07131 amino acid transpo
10	1163	46.2	284	2	T09840 amino acid transpo
11	1153.5	45.8	405	2	A57479 amino acid transpo
12	1142	45.3	386	2	B57479 amino acid transpo
13	1109.5	44.0	411	2	C57479 amino acid transpo
14	966	38.3	284	2	T09843 amino acid transpo
15	557	22.1	441	2	C86378 amino acid transpo
16	507.5	20.1	462	2	T15052 amino acid permeas
17	496	19.7	440	2	P86385 probable lysine an
18	493	19.6	450	2	E96738 hypothetical prote
19	326.5	13.0	442	2	T50687 proline transport
20	323.5	12.8	441	2	T50689 proline transport
21	310	12.3	473	2	T50690 proline transport
22	294.5	11.7	519	2	P96517 hypothetical prote
23	294	11.7	439	2	T50688 proline transport
24	292.5	11.6	436	2	D84782 probable proline t
25	287	11.4	442	2	T50692 proline transport
26	281	11.2	439	2	T47713 amino acid transpo
27	241	9.6	473	2	T04965 probable membrane
28	208	8.3	713	2	S58251 neutral amino acid
29	207.5	8.2	470	2	S47892

30	202	8.0	436	2	T05653 amino acid transpo
31	200	7.9	516	2	T48238 hypothetical prote
32	189.5	7.5	692	2	S37976 hypothetical prote
33	181.5	7.2	484	2	T34016 hypothetical prote
34	176	7.0	494	2	T16658 hypothetical prote
35	176	7.0	543	2	T48239 hypothetical prote
36	175	6.9	423	2	T49959 hypothetical prote
37	169.5	6.7	485	2	T01244 hypothetical prote
38	167.5	6.6	602	2	S55188 hypothetical prote
39	165.5	6.6	486	2	T42254 amino acid permeas
40	165	6.6	462	2	S42372 hypothetical prote
41	158	6.3	571	2	T06737 hypothetical prote
42	157.5	6.3	529	2	E84813 hypothetical prote
43	152.5	6.1	890	2	T21000 hypothetical prote
44	151.5	6.0	449	2	H88022 protein T27A1.5 [1
45	150.5	6.0	434	2	T15799 hypothetical prote

ALIGNMENTS

RESULT 1
A48187
amino acid transport protein I - Arabidopsis thaliana
N:Alternate names: amino acid permease; neutral amino acid carrier
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 27-Jun-1994 #sequence, revision 27-Jun-1994 #text_change 20-Aug-1999
C:Accession: A48187; A47738; S31970
R:Frommer, W.B.; Hummel, S.; Riesmeier, J.W.
Proc. Natl. Acad. Sci. U.S.A. 90, 5944-5948, 1993
A:Title: Expression cloning in yeast of a cDNA encoding a broad specificity amino acid)
A:Reference number: A48187; MIMD:93317600; PMID:8327465.
A:Accession: A48187
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-485 <PRO>
A:Cross-references: EMBL:X67124; NID:g22640; PIDN:CAA47603.1; PID:g22641
A:Note: authors failed to translate the codon ATG for residue 109 as Met
R:Hsu, L.C.; Chiu, T.J.; Chen, L.; Bush, D.R.
Proc. Natl. Acad. Sci. U.S.A. 90, 7441-7445, 1993
A:Title: Cloning a plant amino acid transporter by functional complementation of a yeast
A:Reference number: A47738; MIMD:93361464; PMID:8356039
A:Accession: A47738
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-485 <HSU>
A:Cross-references: GB:L16240; NID:g404018; PIDN:AAA32726.1; PID:g404019
A:Note: sequence extracted from NCBI backbone (NCBIN:136517, NCBI:P:136518)
C:Superfamily: Arabidopsis amino acid transport protein I
C:Keywords: amino acid transport; transmembrane protein; transport protein

Query Match 100.0%; Score 2519; DB 2; Length 485;
Best Local Similarity 100.0%; Pred. No. 2.6e-196;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSFTEGHNHSTASGDAYSDPTKNDDEDERRTGWTMTASAHITTAIGSGVSL 60
DB 1 MMSFTEGHNHSTASGDAYSDPTKNDDEDERRTGWTMTASAHITTAIGSGVSL 60
QY 61 AWAIAQLGMIAGTSLILIFSRITFTSTMLADCYRAPDPVTGRNVTYMDVRSYAGRK 120
DB 61 AWAIAQLGMIAGTSLILIFSRITFTSTMLADCYRAPDPVTGRNVTYMDVRSYAGRK 120
QY 121 VOLGCAVAYGMLIVTGYTTTASISLVANCKNCHDGHADCTISVPMAYEGITIO 180
DB 61 AWAIAQLGMIAGTSLILIFSRITFTSTMLADCYRAPDPVTGRNVTYMDVRSYAGRK 120
QY 121 VOLGCAVAYGMLIVTGYTTTASISLVANCKNCHDGHADCTISVPMAYEGITIO 180
DB 121 VOLGCAVAYGMLIVTGYTTTASISLVANCKNCHDGHADCTISVPMAYEGITIO 180
QY 181 VILSGIPFHKLSFLSIAAVMSFTYATIGLAITVAGKVGKTSMTGAVGVDTAA 240
DB 181 VILSGIPFHKLSFLSIAAVMSFTYATIGLAITVAGKVGKTSMTGAVGVDTAA 240
QY 241 QKWRSPQAVDIAFAVAAYTVLIEIQDTRSSPAENKAMKQASLVGVSTTFPIILGCG 300

Db 241 OKIWRSPQAVGDI AFAYATVLIIEQDTLRSSPAENKAKRASLVGVSTTTFFYLILCGC 300
QY 301 IGYAAGFNAPGDLTDGFFPEFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRN 360
Db 301 IGYAAGFNAPGDLTDGFFPEFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRN 360
QY 361 YPDNKFITSEYVNVPLGKFNISLRLVWRTAYVITTVVAMIPFPFNAILGIAASF 420
Db 361 YPDNKFITSEYVNVPLGKFNISLRLVWRTAYVITTVVAMIPFPFNAILGIAASF 420
QY 421 WPLTVYPPVEMHIAQTKIKKYSARWIALKTMCVCLIVSLAAAGSIAGLISSVKYKPF 480
Db 421 WPLTVYPPVEMHIAQTKIKKYSARWIALKTMCVCLIVSLAAAGSIAGLISSVKYKPF 480
QY 481 RTMHE 485
Db 481 RTMHE 485
RESULT 2
T00620
probable amino acid transport protein T2711.3 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 22-Oct-1999
C:Accession: T00620
R:Federal, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Araujo,
V.; Votokai, V.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.
submitted to the EMBL data library, September 1998
A:Reference number: Z14193
A:Accession: T00620
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-475 <FED>
C:Cross-references: EMBL:AC004122; NID:g3176693; PIDN:AAC34329.1; PID:g3540179; GSPDB:GN
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP-T2711.3
A:Map position: 1
A:Introns: 108/1; 139/2; 211/1; 258/3; 334/3
C:Superfamily: Arabidopsis amino acid transport protein I
C:Keywords: amino acid transport; transmembrane protein; transport protein
Query Match 78.4%; Score 1974.5; DB 2; Length 475;
Best Local Similarity 75.4%; Pred. No. 3.3e-152;
Matches 364; Conservative 58; Mismatches 52; Indels 9; Gaps 2;
QY 1 MKSPNTEGHNHSTAESGDATVSDPTKNVDEGKREKRTGTWLTASAHIIITAVIGSGVLSL 60
Db 1 MDAYN----NPSAVESGDA----AVKSVDDDDGKREKRTGTWLTASAHIIITAVIGSGVLSL 51
QY 61 AWAIAQLGAGTASILLIFSYFTSTMLADCYRAPDPVTGKRNVTYMDVRSYLGGRK 120
Db 52 AWAIAQLGAGTATVLAFAIITYTSTLLADCYRSPDSITGTRNNTYMDVRSYLGGRK 111
QY 121 VOLCGVAQYGNLIGVTGYTTTASISLVAVGKSNCFHDKGTADCTISNYPYMAVFGIIQ 180
Db 112 VOLCGVAQYGNLIGVTGYTTTASISLVAVGKSNCFHDKGTADCTISNYPYMAVFGIIQ 171
QY 181 VILGQINPFIKLSPLSIIAAMVSPFTYATIGIGLAITATVAGGKVGKTSMTGTAAGVDVTA 240
Db 172 ILSQLPNFHLKLSPLSIIAAMVSPFTYATIGIGLAITATVAGGKVGKTSMTGTAAGVDVTA 231
QY 241 OKIWRSPQAVGDI AFAYATVLIIEQDTLRSSPAENKAKRASLVGVSTTTFFYLILCGC 300
Db 232 EKWKLFQATIGDIAFSAFTLILIEQDTLRSSPPENKVMKASLVGVSTTTFFYLILCGC 291
QY 301 IGYAAGFNAPGDLTDGFFPEFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRN 360
Db 292 IGYAAGFNAPGDLTDGFFPEFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKNCKX 351
QY 361 YPDNKFITSEYVNVPLGKFNISLRLVWRTAYVITTVVAMIPFPFNAILGIAASF 420
Db 351 YPDNKFITSEYVNVPLGKFNISLRLVWRTAYVITTVVAMIPFPFNAILGIAASF 420

Db 352 WPOSNFINKSEYSSKVELLQKCRVNLFRVLWRTCYVVLVTTFFVAMIPFPFNAILGLGAF 411
QY 421 WPLTVYPPVEMHIAQTKIKKYSARWIALKTMCVCLIVSLAAAGSIAGLISSVKYKPF 480
Db 412 WPLTVYPPVAMHIAQAKVKKYSRWLNLNLLVLCIVSALAAGSIIGLINSVKYKPF 471
QY 481 RTM 483
Db 472 KNL 474
RESULT 3
T50691
amino acid permease 6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 18-Jul-2001
C:Accession: T50691
R:Rentsch, D.; Hirner, B.; Schmelzer, E.; Frommer, W.B.
Plant Cell 8, 1437-1446, 1996
A:Title: Salt stress-induced proline transporters and salt stress-repressed broad specifi
A:Reference number: Z07208; MUID:96373203; PMID:8776904
A:Accession: T50691
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-481 <REN>
C:Cross-references: EMBL:X95736; PIDN:CAA65051.1
C:Genetics:
A:Gene: AAP6
C:Superfamily: Arabidopsis amino acid transport protein I
Query Match 74.3%; Score 1871; DB 2; Length 481;
Best Local Similarity 73.6%; Pred. No. 8.1e-144;
Matches 345; Conservative 60; Mismatches 62; Indels 2; Gaps 2;
QY 15 ESGDAYTVSDPTKNVDEGKREKRTGTWLTASAHIIITAVIGSGVLSLAWAIAQLGWIAGTS 74
Db 11 QSPPEHEIGDTNKNFDEGDKRTGTWLTGSAHIIITAVIGSGVLSLAWAIAQLGWIAGTS 70
QY 75 ILLIFSPITFTSTMLADCYRAPDPVTGKRNVTYMDVRSYLGGRKVLGVAQYGNLIG 134
Db 71 VLMAFSPITFTSTMLADCYRSPDPVTGKRNVTYMDVRSYLGGRKVLGVAQYGNLIG 130
QY 135 VTGYTTTASISLVAVGKSNCFHDKGTADCTISNYPYMAVFGIIQVILSQINPFIKLS 194
Db 131 ITGYTTTASISLVAVGKSNCFHDKGTADCTISNYPYMAVFGIIQVILSQINPFIKLS 190
QY 195 LSIWAAMVSYATIGIGLAITATVAGGKVGKTSMTGTAAGVDVTAAGKIRSFQAVGDI 253
Db 191 LSIWAAMVSYATIGIGLAITATVAGGKVGKTSMTGTAAGVDVTAAGKIRSFQAVGDI 250
QY 254 AFAYATVLIIEQDTLRSS-PAENKAMKASLVGVSTTTFFYLILCGCIGYAAFGNAPG 312
Db 251 AFAYATVLIIEQDTLRSS-PAENKAMKASLVGVSTTTFFYLILCGCIGYAAFGNAPG 310
QY 313 DELTDGFEPEFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRYDNKFIITSEYS 372
Db 311 NFLTGFYEPFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRYDNKFIITSEYS 370
QY 373 VNPFLGKFNISLRLVWRTAYVITTVVAMIPFPFNAILGLIAGASFVLPVTFVPEMH 432
Db 371 IHPVCCGDFSINFLVWRTSYVTVAVVAMIPFPFNAILGLIAGASFVLPVTFVPEMH 430
QY 433 IAKTKIKKYSARWIALKTMCVCLIVSLLAAAGSIAGLISSVKYKPF 481
Db 431 IAKTKIKKYSARWIALKTMCVCLIVSLLAAAGSIAGLISSVKYKPF 479
RESULT 4
T07130
probable amino acid transporter protein AAP1 - potato (fragment)
C:Species: Solanum tuberosum (potato)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Jun-2000
C:Accession: T07130

Riwart, M.

Submitted to the EMBL Data Library, November 1996

A:Reference number: Z15940

A:Accession: T07130

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-385 <KMA>

A:Cross-references: EMBL:Y09825; PIDN:CAA70966.1

A:Experimental source: cv. Desiree; tuber

C:Genetics:

A:Gene: AAP1

C:Superfamily: Arabidopsis amino acid transport protein 1

C:Keywords: amino acid transport; transmembrane protein; transport protein

Query Match 63.1%; Score 1589.5; DB 2; Length 385;

Best Local Similarity 74.5%; Pred. No. 4e-121;

Matches 287; Conservative 55; Mismatches 42; Indels 1; Gaps 1;

Qy 47 HITAVGSGVLSLMAIAIOLGMIAGTSLILFSPITFTSTMLADCTRAPDPVTKRNY 106

Db 1 HITAVIGSGVLSLMAIAIOLGMIAGTSLILFSPITFTSTMLADCTRAPDPVTKRNY 60

Qy 107 TYMDVNSVYIGRKHVQCGVAQGNLIGTVGTTTASISLVAVGKSNCFDGHADCT 166

Db 61 TYMDVNSHIGVGRKHVQCGVAQGNLIGTVGTTTASISLVAVGKSNCFDGHADCT 120

Qy 167 ISNTPYNAVFGIIVLISQIPNFHKLSELSIMAAVNSFTYATIGLAIATVAG-GKVGK 225

Db 121 IESYPIVITAVIQLVLSQIPNFHKLSELSIMAAVNSFTYATIGLAIATVAG-GKVGK 180

Qy 226 TSMGTAVGVGVVPAQKIRMSFOAVGDIAPFAVAVTVLIEIDTLRSSPANKMKRSL 285

Db 181 TALGVVAVGVGVVPAQKIRMSFOAVGDIAPFAVAVTVLIEIDTLRSSPANKMKRSL 240

Qy 286 VGVSTTTFTTILGCGTGYAFAFNARGLDTPGFPEPMLIDFANACIAVHLIGAVQVF 345

Db 241 AGVSTTTFTTILGCGTGYAFAFNARGLDTPGFPEPMLIDFANACIAVHLIGAVQVF 300

Qy 346 AQPFOVEKKCNRYPDNKEITSEYVNVPELCKFNISLFLVWRTAVVITVVMVIF 405

Db 301 QPFGYVBERGSRPDSKFIITSEYAMQVPGTNNLFLVWRTAVVITVVMVIF 360

Qy 406 PFRVAILGLIGASFWPLTYFPVE 430

Db 361 PFRVAILGLIGASFWPLTYFPVE 385

RESULT 5

T10100

amino acid transport protein AAP1 - caenor bean

N:Alternate names: amino acid permease; amino acid transporter

C:Species: Ricinus communis (castor bean)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T10100

R:Marvier, A.C.; Neelam, A.; Bick, J.A.; Hall, J.L.; Williams, L.E.

Biochim. Biophys. Acta 1373, 321-331, 1998

A:Title: Cloning of a cDNA coding for an amino acid carrier from Ricinus communis (RCAAF

A:Reference number: Z16947; MIDB:98405925; PMID:9733991

A:Accession: T10100

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-486 <MAR>

A:Cross-references: EMBL:AJ007574; NID:G3293030; PIDN:CAA07563.1; PID:G3293031

A:Experimental source: cultivar sanguineous

C:Genetics:

A:Gene: AAP1

C:Function: transport of basic amino acids e.g. histidine

A:Superfamily: Arabidopsis amino acid transport protein I

C:Keywords: amino acid transport; transmembrane protein; transport protein

Query Match 61.2%; Score 1540.5; DB 2; Length 486;

Best Local Similarity 58.2%; Pred. No. 4.9e-117;

Matches 285; Conservative 82; Mismatches 106; Indels 17; Gaps 2;

Qy 5 NTEGHNRSTAESGDATVYSDPKN-----VDEGRKRRTGTWLTASAHITVIG 54

Db 4 NTAANKH-----PHQVDPVDSINMOTQVGSKMLDDGKRRTGTWLTASAHITVIG 56

Qy 55 SGVLSLMAIAIOLGMIAGTSLILFSPITFTSTMLADCTRAPDPVTKRNYTMDVRS 114

Db 57 SGVLSLMAIAIOLGMIAGTSLILFSPITFTSTMLADCTRAPDPVTKRNYTMDVRS 116

Qy 115 YLGRKTYQLCGVAQGNLIGTVGTTTASISLVAVGKSNCFDGHADCTISNYMA 174

Db 117 NLGAKAKVLCGFQVYNLGVNIGVATISISMAIKSNCFHSGGKRPCHIANPYMT 176

Qy 175 VFGIIVLISQIPNFHKLSELSIMAAVNSFTYATIGLAIATVAGKVGKSTMTAVG 234

Db 177 AFGIAELIFQIDDPQQLMWLSTLAIVNSFTYATIGLAIATVAGKVGKSTMTAVG 236

Qy 235 VDTVAQKIRMSFOAVGDIAPFAVAVTVLIEIDTLRSSPANKMKRSLVGVSTTFP 294

Db 237 ANVPTQKIRMSFOAVGDIAPFAVAVTVLIEIDTLRSSPANKMKRSLVGVSTTFP 296

Qy 295 YLGGCIGVAFGNAPGFLDFGFPEPMLIDFANACIAVHLIGAVQVFAQPIFOVE 354

Db 297 YMLGCGTGYAFAFNARGLDTPGFPEPMLIDFANACIAVHLIGAVQVFAQPIFOVE 356

Qy 355 KCKNRNPDNKEITSEYVNVPELCKFNISLFLVWRTAVVITVVMVIFPFRNALGL 414

Db 357 KAAQRPVDSGFTTKIKIPVDFPRFNLFRSQRFLFVFTVISMILPFRNDIVGL 416

Qy 415 IGAASFPLTYFPVEVEMHIAQTKIKYASRWIALTMCVCVILSLAAAGSISGLISSV 474

Db 417 LGAGFPLTYFPVEVEMHIAQTKIKYASRWIALTMCVCVILSLAAAGSISGLISSV 476

Qy 475 KTYPEFRTH 484

Db 477 KSVKRFQISY 486

RESULT 6

H96802

probable amino acid carrier [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: H96802

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzilli,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: AB6141; MIDB:21016719; PMID:11130712

A:Accession: H96802

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-476 <STO>

A:Cross-references: GB:AE005173; NID:G11079491; PIDN:AA629203.1; GSPDB:GNO0141

C:Genetics:

A:Gene: F2P24.9

A:Map position: 1

C:Superfamily: Arabidopsis amino acid transport protein I

Query Match 59.7%; Score 1504; DB 2; Length 476;

Best Local Similarity 57.6%; Pred. No. 4.3e-114;

Matches 273; Conservative 97; Mismatches 100; Indels 4; Gaps 4;

Qy 10 NHSTASGDAVYSDPKNVDEGRKRRTGTWLTASAHITVIGSGVLSLMAIAIOLG 69

Db 4 NHQTVLAVD-MPQGSKYLDDGKMKRTGSWTSAHITVIGSGVLSLMAIAIOLG 62

Db 197 GLGLGVSKVENNEIGSLTGVTVLTSLGTVTSQKIMRTPQSLGNIAFAVSIMLI 256
Qy 265 EIDDTLRSSAEKAKMKRSLVGVSTTFPYLLCCGIGYAAFGNNAQDPLTDGPFEPF 324
Db 257 EIDDTVSPAEAVETMRKATFVSAVATTFYMLCCGVGYAAGNAPENLTAHQGFNPY 316
Qy 325 MLIDPANACIAVHLIGAYVQAPDPIQFQVEKKCNRYNDNFITSESVNVPFLGK-FNI 383
Db 317 MLIDIANLAIHVLGAYVQCPPLFAFVEKASRPFSESEFVKEIKIQI-PPGRFNL 375
Qy 384 SLPLRWRTAVVITTVAMIFPPFPAIIGLGAASNPPLVYPPVEMHIAQTIKKYS 443
Db 376 NLPLRWRTAVVITTVAMIFPPFPAIIGLGAASNPPLVYPPVEMHIAQTIKKYS 435
Qy 444 RMIALKTMCCVCHIVSLAAGSIAGLISSVTVKPRFT 482
Db 436 KMVCLQVLSVTCFVGSVAAAGSVIGVSDLVKVPQS 474

RESULT 9

T07131
C:Species: Solanum tuberosum (potato)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Jun-2000
C:Accession: T07131
R:Kwart, M.
Submitted to the EMBL Data Library, November 1996
A:Reference number: Z15940
A:Accession: T07131
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-376 <KMA>
A:Cross-references: EMBL:Y09826; PIDD:CAA70969.1
A:Experimental source: cv. Desiree, leaf
C:Genetics:
A:Gene: AAP2
C:Superfamily: Arabidopsis amino acid transport protein I
C:Keywords: amino acid transport; transmembrane protein; transport protein

Query Match 52.6%; Score 1326; DB 2; Length 376;

Best Local Similarity 62.9%; Pred. No. 8.7e-100;
Matches 237; Conservative 68; Mismatches 70; Indels 2; Gaps 2;

Qy 47 HIITAVIGSVLISLANIAQLGNIAGTSLILFSTFTYFTSTMLADCYRADPVTGKKNY 106
Db 1 HIITAVIGSVLISLANIAQLGNIAGTSLILFSTFTYFTSTMLADCYRADPVTGKKNY 60
Qy 107 TYMDVRSYLGKRVOLCGVAYQGNLIGVTVGTTTASISLVAVGKSNCFHDKGHTADCT 166
Db 61 TYMDAVANLGGQVKICVITQANLFGVAIGTTTASISLVAVGKSNCFHDKGHTADCT 120
Qy 167 ISNYPYAVFGIIVLISQIPNFKLSPLSMAAVSFTYATIGLAIAVTA-GKRVGK 225
Db 121 VSSPYMIFGVWEIIFSQIPDQISMLISIAVAWSFTYIGLGVAVAEIGKI-E 179
Qy 226 TSMGTAVGVDTAAQKIRWSPQAVDIAFAVAATVLEIIDDTRSRPAENKMKRSL 285
Db 180 GSLGIGISGIEVEMOKIRWSPQALCAIFAYSYSLILEIIDDTRSRPAENKMKRSL 239
Qy 286 VGVSTTFPYLLCCGIGYAAFGNNAQDPLTDGPFEPFMLIDPANACIAVHLIGAYV 345
Db 240 ISVAVTVFVFMGLCGCFGYAAGFQSPGNILTGFGFNPMLIDIANVAIVHLGAYV 299
Qy 346 AQPFFQVEKKCNRYNDNFITSESVNVPFLGKFNISLPLRWRTAVVITTVAMIF 405
Db 300 CQPLFAVETKATFVSAVATTFYMLCCGVGYAAGNAPENLTAHQGFNPY 359
Qy 406 PFPNAILGLGAASNP 422
Db 360 PFPNDVVGIIAGAFQFMP 376

RESULT 10

T09840
amino acid transport protein AAP2 - castor bean (fragment)
N:Alternate names: amino acid permease; amino acid transporter
C:Species: Ricinus communis (castor bean)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09840
R:Black, J.A.; Neelam, A.; Hall, J.L.; Williams, L.E.
Plant Mol. Biol. 36, 377-385, 1998
A:Title: Amino acid carriers of Ricinus communis expressed during seedling development:
A:Reference number: Z16878; MUID:98145472; PMID:9484478
A:Accession: T09840
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-284 <BIG>
A:Cross-references: EMBL:Y11121; NID:G1839021; PIDD:CAA72006.1; PID:G1839022
A:Experimental source: strain Sanguineus; tissue-type root; clone RCAA2
C:Superfamily: Arabidopsis amino acid transport protein I
C:Keywords: amino acid transport; transmembrane protein; transport protein

Query Match 46.2%; Score 1163; DB 2; Length 284;

Best Local Similarity 75.0%; Pred. No. 1e-86;
Matches 213; Conservative 37; Mismatches 34; Indels 0; Gaps 0;

Qy 63 AIALGMIAGTSLILFSTFTYFTSTMLADCYRADPVTGKKNYMDVRSYLGKRVQ 122
Db 1 AIALGMIAGTSLILFSTFTYFTSTMLADCYRADPVTGKKNYMDVRSYLGKRVQ 60
Qy 123 LCGVAYQGNLIGVTVGTTTASISLVAVGKSNCFHDKGHTADCTISNYPYAVFGIIV 182
Db 61 LCGVAYQGNLIGVTVGTTTASISLVAVGKSNCFHDKGHTADCTISNYPYAVFGIIV 120
Qy 183 LSQIPNFKLSPLSMAAVSFTYATIGLAIAVTAAGKQKSMGTAVGVDTAAOK 242
Db 121 LSQIPNFKLSPLSMAAVSFTYATIGLAIAVTAAGKQKSMGTAVGVDTAAOK 180
Qy 243 IWRSPQAVDIAFAVAATVLEIIDDTRSRPAENKMKRSLVGVSTTFPYLLCCGIG 302
Db 181 IWRSPQAVDIAFAVAATVLEIIDDTRSRPAENKMKRSLVGVSTTFPYLLCCGIG 240
Qy 303 YAAFGNNAQDPLTDGPFEPFMLIDPANACIAVHLIGAYVFA 346
Db 241 YAAFGNNAQDPLTDGPFEPFMLIDPANACIAVHLIGAYVFA 284

RESULT 11

A57479

amino acid transport protein AAP3 - Arabidopsis thaliana
N:Alternate names: amino acid permease; amino acid transporter
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Jun-2000
C:Accession: A57479; S51168
R:Fischer, W.N.; Kwart, M.; Hummel, S.; Frommer, W.B.
J. Biol. Chem. 270, 16315-16320, 1995
A:Title: Substrate specificity and expression profile of amino acid transporters (AAPs)
A:Reference number: A57479; MUID:95332342; PMID:7608199
A:Accession: A57479
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-405 <FIS>
A:Cross-references: GB:X77499
R:Fischer, W.N.
Submitted to the EMBL Data Library, January 1994
A:Reference number: S51168
A:Accession: S51168
A:Molecule type: mRNA
A:Residues: 1-21, 'V', 23-60, 'GWLADPVMILFSA', 61-136, 'KSGKGRPCIMNSNP', 137-205, 'IMRTPOLGDI
A:Cross-references: EMBL:X77499; NID:9608668; PID:9608668
C:Genetics:
A:Gene: AAP3
C:Superfamily: Arabidopsis amino acid transport protein I
C:Keywords: amino acid transport; transport protein

Query Match 45.8%; Score 1153.5; DB 2; Length 405;

Db 55 STLLSDCYRTGDPVSGKRNYYMDAVRSILGFRPKICGLIOYUNLFGITVGYTAAASIS 114
 QY 147 LVAVGKSCNFHDKGHTADCTISNYPYMAVFGIIQVILSQIPNFHKLISFLSIAAAMSFTY 206
 Db 115 MWAIKRSCNFH-----YIMFGVTEILLSQIDFQIWWLSIVAAIMSFTY 160
 QY 207 ATIGGLAIATVAGGKVKTSMTGTAVGVDTVAQAKIWRSPQAVGDTAFAYATVLIIEI 266
 Db 161 SAIGLALGLIOVANGVVRGSLTGISGA-VTQTK-----YSYSVLLIEI 205
 QY 267 QDTRSSPAENKAMKRASLVGVSTTTFFYILCGCIGYAAFGNNAAGDGLTDFGFFPPFWL 326
 Db 206 QDTRSPAESKTMKIAIRISIAVTITFVYMLCGCGYAAFADKAPGNLLTGFGFYNPFWL 265
 QY 327 IDPANACTAVHLIGAYQVFAQPIQFQVBEKKNRNPDKNFTTSYSVNVP-FLGKNISL 395
 Db 266 LDVAN-----AQPIAFIEKQLAARPDSDLVTKEYEIRIQFGRSPYKQNV 311
 QY 386 FRLVWRTAYVITTVVAMIEPFFENALGLIGAASEFWLTVFVPEMHIQTKIKKYSARW 445
 Db 312 FRANYRSGFVLTIVISMLPFFNDVVUGILGALGFW-----RKVRNSMKW 357
 QY 446 IALKTMCVCLIVLSLLAAAGSIAGLISSVKTKPPTW 484
 Db 358 VCLQWLSGCLMLITLVAGVSIAGVMDLKYKPKTTY 396

RESULT 13
 C57479
 amino acid transport protein AAPs - Arabidopsis thaliana
 N:Alternate names: amino acid permease; amino acid transporter
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Jun-2000
 C:Accession: C57479; S51170
 R:Fischer, W.N.; Kwart, M.; Hummel, S.; Frommer, W.B.
 J. Biol. Chem. 270, 16315-16320, 1995
 A:Title: Substrate specificity and expression profile of amino acid transp
 A:Reference number: A57479; MUID:95332342; PMID:7608199
 A:Accession: C57479
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-411 <FIS>
 A:Cross-references: GB:X77501
 R:Fischer, W.N.
 submitted to the EMBL Data Library, January 1994
 A:Reference number: S51168
 A:Accession: S51170
 A:Molecule type: mRNA
 A:Residues: 1-58, 'GWGQGVPMLLPSF', 59-88, 'I', 90-134, 'MKNPDPCVHGVN', 135-21
 A:Cross-references: EMBL:X77501; NID:9608672; PIDN:CAA54632.1; PID:9608673
 C:Genetics:
 A:Gene: AAPs
 C:Superfamily: Arabidopsis amino acid transport protein I
 C:Keywords: amino acid transport; transport protein

Query Match 44.0%; Score 1109.5; DB 2; Length 411;
 Best Local Similarity 47.1%; Pred. No. 3,3e-82;
 Matches 216; Conservative 80; Mismatches 86; Indels 77; Gaps

QY 30 DDGREGKTGTLWTASAHITAVIGSGVLSLAWAJAQLGWTAGTISILLIFSTYFTSTM 89
 Db 21 DDGGRPKRTGTWVTASAHITAVIGSGVLSLAWAAQI-----VTFYFTSL 66
 QY 90 LADCYRAPDPVTGKRNYYMDVVRSLGCRKVOLCGVAQVGNLIGTVGVYITIASISLVA 149
 Db 67 LCSCYRSGDSVTGKENTYMDATHSNLGGIKVKVGVVQYVNLFGTALGYTIASISLVA 126
 QY 150 VKGSCNFHDKGHTADCTISNYPYMAVFGIIQVILSQIPNFHKLISFLSIAAAMSFTYATI 209
 Db 127 IORTSC-----QQYNIAFGVIQIIFISQIPDFQOLWMLSIVAAMVSFAYSAL 172
 QY 210 GIGLAIAIVAGGKVKTSMTGTAVG-----DVTAAQIKWRSPQAVGDTAFAYATVLI 264

```

Db 173 GLGIVGKVENNEIKGSLGVTVLSTGVTSQK-----YYSMLI 218
Qy 265 EIQTTLSSPAENKAMKRAASLVGVSTTTFFYLCCIGAAFGNNAFGDPLTDDGFPEPF 324
Db 219 EIQTTLSSPAENKAMKRAASLVGVSTTTFFYLCCIGAAFGNNAFGDPLTDDGFPEPF 278
Qy 325 WLIDPANACIAVHLIGAYQVFAPIQFQVEKKCNPNPNKETSSEYVNPFLGK-FNI 383
Db 279 WLIDPAN-----IQPLFAFVEKESRPRPESEVTEIKIQ-FPGKPFNL 323
Qy 384 SLFLVWRRTAVVITTVVAMIFPPENAIIGLIGASFPMLTYFVEVEMHIAQTIKXISA 443
Db 324 NLFLVWRRTAVVITTVVAMIFPPENAIIGLIGASFPMLTYFVEVEMHIAQTIKXISA 369
Qy 444 RMIALKTKCVCLIVSLAAGSIAGLISSVKTYKPEPT 482
Db 370 KMWCLQVLSVTCIFVSAABAAGSVIGIVSDKRYKPEFS 408

```

RESULT 14

T09843

amino acid transport protein AAP1 - caenor bean (fragment)

N:Alternate names: amino acid permease; amino acid transporter

C:Species: Ricinus communis (castor bean)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T09843

R:Block: J.A.; Neelam, A.; Hall, J.L.; Williams, L.E.

Plant Mol. Biol. 36, 377-385, 1998

A:Title: Amino acid carriers of Ricinus communis expressed during seedling development:

A:Reference number: Z16878; MUID:98145472; PMID:9484478

A:Accession: T09843

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-284 <BIC>

A:Cross-references: EMBL:Z68759, NID:G2995320, PIDD:CAA92992.1, PIDD:G2995321

A:Experimental source: strain Sanguiinas; tissue-type root; clone RCAA1

C:Superfamily: Arabidopsis amino acid transport protein I

C:Keywords: amino acid transport; transmembrane protein; transport protein

Query Match 38.3%; Score 966; DB 2; Length 264;

Best Local Similarity 61.3%; Pred. No. 9.2e-71;

Matches 174; Conservative 50; Mismatches 60; Indels 0; Gaps 0;

```

Qy 63 AIAQIGMIAGTSLILFSTFTYFTSMADCYRAPDVTGKKNYTMVDVNSYSGRRVQ 122
Db 1 AIAQIGMIAGTSLILFSTFTYFTSMADCYRAPDVTGKKNYTMVDVNSYSGRRVQ 60
Qy 123 LCGVAOYGNLIGTVGVTITASISIVAVGKSNCFHDKHTADCTISNYPMAVFGIIQVT 182
Db 61 LCGVQYTNLIFGVNIGTITASISIMPAIKRSNCFHSGGKAPCHINANPMIAFGIAIIT 120
Qy 183 LSGIPIHFHKLSPFISIMAAVNSFTYATIGIALATVAGGKVGKTSMTGTAVGVDTAAQK 242
Db 121 FSGIPIHFHKLSPFISIMAAVNSFTYATIGIALATVAGGKVGKTSMTGTAVGVDTAAQK 180
Qy 243 IWRSPQAVGDIAPVAYATVLIETIOPTLSSPAENKAMKRAASLVGVSTTTFFYLCCIG 302
Db 181 IWRSPQAVGDIAPVAYATVLIETIOPTLSSPAENKAMKRAASLVGVSTTTFFYLCCIG 240
Qy 303 YAAFGNNAFGDPLTDDGFPEPFWLIDPANACIAVHLIGAYQVFA 346
Db 241 YAAFGNNAFGDPLTDDGFPEPFWLIDPANACIAVHLIGAYQVFA 284

```

RESULT 15

C86378

protein F21J9.6 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C86378

R:Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

```

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maizel, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A66141; MUID:21016719; PMID:11130712
A:Accession: C86378
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-441 <STO>
A:Cross-references: GB:AE005172; NID:G9743356; PIDD:AAF97980.1; GSPDB:GN00141
C:Genetics:
A:Gene: F21J9.6
A:Map position: 1
C:Superfamily: Arabidopsis amino acid transport protein I

```

Query Match 22.1%; Score 557; DB 2; Length 441;

Best Local Similarity 31.1%; Pred. No. 2.1e-37;

Matches 145; Conservative 80; Mismatches 187; Indels 54; Gaps 12;

```

Qy 27 KNVDE--DGEKRTGTWLTSAHIIIPAVIGSVLSLAWAIAQIGMIAGTSLILFSEFTY 84
Db 17 KNVDMLPIITSSNAKMWASAFINVTAVYAGVLSLIPAMSLGMPGVITIMWSMITL 76
Qy 85 FTSMADCYRAPDVTGKKNYTMVDVNSYSGRRVQCGYQYQNU--IGTVGTYIT 142
Db 77 YTLQWQEMH---ETVPGKRLRYHE--LQGHAFGKELGLMTVPPQQLVBEVGVIYVWT 132
Qy 143 ASISLVAVGKSNCFHDKHTADCTISNYPMAVFGIIQVLSQIPIHFHKLSPFISIMAAV 201
Db 133 GGAISLKVHQLVC-----PDCKELRTTFWIMIRASVHFVISHLPNNSISISLAAV 185
Qy 202 MSFTYATIGIALATVAGGKVGKTSMTGTAVGVDTAAQKIWRSPQAVGDIAPVAYAT 261
Db 186 MSFTYATIGIALATVAGGKVGKTSMTGTAVGVDTAAQKIWRSPQAVGDIAPVAYAT 236
Qy 262 VLIETIOPTLSSPAENKAMKRAASLVGVSTTTFFYLCCIGIYVAFGNNAFGDPLDFFG 319
Db 239 VLIETIOPTLSSPAENKAMKRAASLVGVSTTTFFYLCCIGIYVAFGNNAFGDPLDFFG 319
Qy 319 FPEPFWLIDPANACIAVHLIGAYQVFAPIQFQVE---KCNPNPNKETSSEYVNPFLGK 375
Db 299 FPEPFWLIDPANACIAVHLIGAYQVFAPIQFQVE---KCNPNPNKETSSEYVNPFLGK 339
Qy 376 FPLGKPNISL-FRLVWRRTAVVITTVVAMIFPPENAIIGLIGASFPMLTYFVEVEMHIA 434
Db 340 FPLGKPNISL-FRLVWRRTAVVITTVVAMIFPPENAIIGLIGASFPMLTYFVEVEMHIA 394
Qy 435 QTKIKKYSARMIALKTWCYCLIVSLAAGSIAGLISSVKTYKPE 480
Db 395 LKKKRRFGLSWTAMWFCIIIGVLTILADIGLRTIIINAKTYKPE 440

```

Search completed: November 22, 2002, 13:30:45

Job time : 23 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2002, 13:28:57 ; Search time 35 Seconds
(without alignments)
2955.224 Million/cell updates/sec

Title: US-09-854-562-2

Sequence: 1 MSFNTGHNHSTAESGDY.....SIAGLISSVTKYKPRFTMHE 485

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

SPTREMBL_21:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp Vertebrate:*

14: sp_undefined:*

15: sp_virus:*

16: sp_bacteriophage:*

17: sp_archaeophages:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2519	100.0	485	10	Q42400 arabidopsis
2	2510	99.6	485	10	Q48546 arabidopsis
3	1974.5	78.4	475	10	Q80592 arabidopsis
4	1871	74.3	481	10	P92934 arabidopsis
5	1589.5	63.1	385	10	P93561 solanum tub
6	1540.5	61.2	486	10	Q82044 ricinus com
7	1504	59.7	476	10	Q39134 arabidopsis
8	1501	59.6	493	10	Q38967 arabidopsis
9	1496	59.4	466	10	Q8RW48 arabidopsis
10	1495	59.3	466	10	Q9FN04 arabidopsis
11	1494	59.3	466	10	Q39135 arabidopsis
12	1484.5	58.9	379	10	Q9ZPM7 nepenthes a
13	1471.5	58.4	488	10	Q8RZP7 oryza sativ
14	1462	58.0	481	10	Q93X13 vicia faba
15	1442.5	57.3	466	10	Q9ZRS1 ricinus com
16	1441.5	57.2	513	10	Q9ARG2 glycine max

17	1427.5	56.7	480	10	Q39136 arabidopsis
18	1427.5	56.7	509	10	Q9ZRE2 vicia faba
19	1417.5	56.3	475	10	Q93X15 vicia faba
20	1406.5	55.8	476	10	Q9C612 arabidopsis
21	1326	52.6	376	10	P93562 solanum tub
22	1294.5	51.4	377	10	Q9ZPM6 nepenthes a
23	1285	51.0	486	10	Q93X14 vicia faba
24	1200.5	47.7	376	10	Q9ZPM5 nepenthes a
25	1198.5	47.6	469	10	Q83M9 oryza sativ
26	1179	46.8	467	10	Q9FP99 arabidopsis
27	1163	46.2	284	10	P93506 ricinus com
28	1016.5	40.4	263	10	Q9SEW7 vicia faba
29	966	38.3	284	10	Q41132 ricinus com
30	860.5	34.2	261	10	Q9SEW6 vicia faba
31	840.5	33.4	303	10	Q941E6 arabidopsis
32	744.5	29.6	259	10	Q9SEW5 vicia faba
33	557	22.1	441	10	Q9LRB5 arabidopsis
34	548	21.8	441	10	Q9SR44 arabidopsis
35	518	20.6	154	10	Q8VMT1 narcissus p
36	514	20.4	161	10	Q8VMT0 narcissus p
37	507.5	20.1	462	10	Q40414 nicotiana s
38	501	19.9	446	10	Q9FKS8 arabidopsis
39	496	19.7	440	10	Q9C6M2 arabidopsis
40	493	19.6	450	10	Q9M9H9 arabidopsis
41	485	19.3	809	10	Q9C7J3 arabidopsis
42	461	18.3	424	10	Q9C9J0 arabidopsis
43	437	17.3	446	10	Q24405 arabidopsis
44	436.5	17.3	455	10	Q9SGJ3 arabidopsis
45	436.5	17.3	479	10	Q9SS86 arabidopsis

ALIGNMENTS

RESULT 1

Q42400 PRELIMINARY; PRT; 485 AA.

AC Q42400; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Amino acid permease 1 (P19C14.3 protein).

GN AAP1 OR P19C14.3 OR ZCF54.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliophyta; Magnoliopsida; core eudicots; Rosidae; eucosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SEF. LANDSBERG;

RX MEDLINE=93317600; PubMed=8327465;

RA Frommer W.B., Hummel S., Riesmeier J.W.;

RT "Expression cloning in yeast of a cDNA encoding a broad specificity amino acid permease from Arabidopsis thaliana."

RL Proc. Natl. Acad. Sci. U.S.A. 90:5944-5948(1993).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=93361464; PubMed=8356039;

RA Heu L.C., Chou T.U., Chen U., Bush D.R.;

RT "Cloning a plant amino acid transporter by functional complementation of a yeast amino acid transport mutant."

RL Proc. Natl. Acad. Sci. U.S.A. 90:7441-7445(1993).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Liu S.X., Sakano H., Yu G., Lee J.M., Lenz C., Pham P., Tortum M., Chin C., Chou J., Choi E., Chung M., Gonzalez A., Hwang B., Liu A., Vayenberg M., Alatai H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;

RT "The sequence of BAC P19C14 from Arabidopsis thaliana chromosome 1."

Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RL [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CV, COLUMBIA;
 RA Kato A., Kameda Y.;
 RT "Long repeat sequence within a genomic region located around the 100
 map unit of chromosome 1 in Arabidopsis thaliana.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=20018182; PubMed=10548732;
 RA Kato A., Suzuki M., Kuwahara A., Ooe H., H-Inaba K., Kameda Y.;
 RT "Isolation and analysis of a cDNA within a 300 kb Arabidopsis thaliana
 genomic region located around the 100 map unit of chromosome 1.";
 RL Gene 239:309-316(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=20263341; PubMed=10805591;
 RA Suzuki M., Kato A., Kameda Y.;
 RT "An RNA-Binding protein, AtBBP1, is expressed in actively
 proliferative regions in Arabidopsis thaliana.";
 RL Plant Cell Physiol. 41:282-288(2000).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RA Suzuki M., Kato A., Nagata N., Kameda Y.;
 RT "An xylanase, AtXyl1, is predominantly expressed in vascular bundles,
 RT and four putative xylanase genes were identified in the Arabidopsis
 thaliana genome.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X67124; CAA47603.1; -;
 DR EMBL; L16240; AAA32726.1; -;
 DR EMBL; AC008051; AAF82252.1; -;
 DR EMBL; AB077822; BAB83868.1; -;
 DR InterPro; IPR002422; AA/re1_prtase2.
 DR InterPro; IPR001395; Aldo/ket_red.
 DR Pfam; PF01490; Aa_trans; 1
 DR PROSITE; PS00063; ALDOKETO_REDUCTASE_3; UNKNOWN 1.
 SQ SEQUENCE 485 AA; 52895 MW; 722E3895937CCC5B CRC64;

Query Match 100.0%; Score 2519; DB 10; Length 485;
 Best Local Similarity 100.0%; Pred. No. 2.2e-188;
 Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKEPTEGHNHSTAESGDAYTVSDPTKNVDEGDKREKRTGTWLTASAHITAVIGSVLSL 60
 DB 1 MKEPTEGHNHSTAESGDAYTVSDPTKNVDEGDKREKRTGTWLTASAHITAVIGSVLSL 60
 QY 61 AWATAQLGWIAGTSSILLIPSFITYFTSTMLADCYRAPDPVTGKNTYTMVVRVYLGRK 120
 DB 61 AWATAQLGWIAGTSSILLIPSFITYFTSTMLADCYRAPDPVTGKNTYTMVVRVYLGRK 120
 QY 121 VOLCGVAQYGNLIGVTGYTTTASISLVAGKSNCFHDKGHTADCTISNYPYMAVFGIIQ 180
 DB 121 VOLCGVAQYGNLIGVTGYTTTASISLVAGKSNCFHDKGHTADCTISNYPYMAVFGIIQ 180
 QY 181 VILSQIPNFHKLFLSILMAAVMSFTYATIGIGLAITATVAGGKVGKTSMTGTAVGVDVTA 240
 DB 181 VILSQIPNFHKLFLSILMAAVMSFTYATIGIGLAITATVAGGKVGKTSMTGTAVGVDVTA 240
 QY 241 QKIWRSPQAVGDIAPAYATVLEIQTDLRSSPAENKAMKASLVGVSTTTFFPILCGC 300
 DB 241 QKIWRSPQAVGDIAPAYATVLEIQTDLRSSPAENKAMKASLVGVSTTTFFPILCGC 300
 QY 301 IGYAAGFNAPGDELTDGFFEPFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRN 360
 DB 301 IGYAAGFNAPGDELTDGFFEPFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRN 360
 QY 361 YPDNKFITSEYVNVPLGKFNLSLRLVWRVAYVITTVAMIPFFNALIGLIGAS 420
 DB 361 YPDNKFITSEYVNVPLGKFNLSLRLVWRVAYVITTVAMIPFFNALIGLIGAS 420

QY 421 WPLTVYFPVEMHIAQTKIKKYSARWIALKTMCTVCLIVSLAAAGSIAGLISSVKTYPF 480
 DB 421 WPLTVYFPVEMHIAQTKIKKYSARWIALKTMCTVCLIVSLAAAGSIAGLISSVKTYPF 480
 QY 481 RTMHE 485
 DB 481 RTMHE 485

RESULT 2

Q48546
 ID 048546 PRELIMINARY; PRT; 485 AA.
 AC 048546;
 DT 01-JUN-1998 (T-EMBLrel. 06, Created)
 DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Neutral amino acid transport system II.
 GN NAT2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=93361464; PubMed=8356039;
 RA Hsu L.-C., Chiot T.-J., Chen L., Bush D.R.;
 RT "Cloning a plant amino acid transporter by functional complementation
 of a yeast amino acid transport mutant";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7441-7445(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RA Chang H.-C., Bush D.R.;
 RT "Genomic Sequence of NAT2/AAPI of Arabidopsis";
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF031649; BAB87674.1; -;
 DR InterPro; IPR002422; AA/re1_prtase2.
 DR InterPro; IPR001395; Aldo/ket_red.
 DR Pfam; PF01490; Aa_trans; 1.
 DR PROSITE; PS00063; ALDOKETO_REDUCTASE_3; UNKNOWN 1.
 SQ SEQUENCE 485 AA; 52943 MW; C4F3C49A538FC06B CRC64;

Query Match 99.6%; Score 2510; DB 10; Length 485;
 Best Local Similarity 99.8%; Pred. No. 1.1e-187;
 Matches 484; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKEPTEGHNHSTAESGDAYTVSDPTKNVDEGDKREKRTGTWLTASAHITAVIGSVLSL 60
 DB 1 MKEPTEGHNHSTAESGDAYTVSDPTKNVDEGDKREKRTGTWLTASAHITAVIGSVLSL 60
 QY 61 AWATAQLGWIAGTSSILLIPSFITYFTSTMLADCYRAPDPVTGKNTYTMVVRVYLGRK 120
 DB 61 AWATAQLGWIAGTSSILLIPSFITYFTSTMLADCYRAPDPVTGKNTYTMVVRVYLGRK 120
 QY 121 VOLCGVAQYGNLIGVTGYTTTASISLVAGKSNCFHDKGHTADCTISNYPYMAVFGIIQ 180
 DB 121 VOLCGVAQYGNLIGVTGYTTTASISLVAGKSNCFHDKGHTADCTISNYPYMAVFGIIQ 180
 QY 181 VILSQIPNFHKLFLSILMAAVMSFTYATIGIGLAITATVAGGKVGKTSMTGTAVGVDVTA 240
 DB 181 VILSQIPNFHKLFLSILMAAVMSFTYATIGIGLAITATVAGGKVGKTSMTGTAVGVDVTA 240
 QY 241 QKIWRSPQAVGDIAPAYATVLEIQTDLRSSPAENKAMKASLVGVSTTTFFPILCGC 300
 DB 241 QKIWRSPQAVGDIAPAYATVLEIQTDLRSSPAENKAMKASLVGVSTTTFFPILCGC 300
 QY 301 IGYAAGFNAPGDELTDGFFEPFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRN 360
 DB 301 IGYAAGFNAPGDELTDGFFEPFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRN 360

QY	361	YPKPKFTSSYSNVNPEFLKRNISLPELWMPRAVYVITTVAMLPFPFNALIGLIGASAF	420
Db	361	YPDKKFTTSSYSNVNPEFLKRNISLPELWMPRAVYVITTVAMLPFPFNALIGLIGASAF	420
QY	421	WPLTVPEPVMHIAOTKIKKYSARMIATKMCYVCLIVSLAAGSIAGLISSVKTYKPF	480
Db	421	WPLTVPEPVMHIAOTKIKKYSARMIATKMCYVCLIVSLAAGSIAGLISSVKTYKPF	480
QY	481	RTMHE 485	
Db	481	RTMHE 485	
RESULT 3			
ID	080592	PRELIMINARY;	PRT; 475 AA.
AC	080592;		
DT	01-NOV-1998	(TRENBLREL. 08, Created)	
DT	01-NOV-1998	(TRENBLREL. 08, Last sequence update)	
DT	01-MAR-2002	(TRENBLREL. 20, Last annotation update)	
DE		Putative amino acid permease.	
GN	12711.3.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eumustos II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_TaxID=3702;		
FN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. COLUMBIA;		
RA	Federstepel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,		
RA	Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,		
RA	Oji O., Osborne B.I., Shim P., Sun H., Tortum M., Vysotskaya V.S.,		
RA	Yu G., Becker J., Theologis A., Davis R.W.;		
RL	Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; ACC04122; AAC34329.1; -		
DR	InterPro: IPR002422; AA/rel_primease2.		
DR	Plam; PFI01490; Aa_trans; 1.		
SQ	SEQUENCE 475 AA; 51815 MW; B6DBE571E2FE7P92 CRC64;		
Query Match 78.4%; Score 1974.5; DB 10; Length 475;			
Best Local Similarity 75.4%; Pred. No. 6.4e-146;			
Matches 364; Conservative 58; Mismatches 52; Indels 9; Gaps 2.			
QY	1	MKSFNTGHNHSTAESGDVAYVSDPTKVNDEDEKRTGTWLTAHIIITAVIGSGVLSI	60
Db	1	MDAYN----NPSAVESGDA----AVKSYDDGGRKRGTFWTASAHIIITAVIGSGVLSI	51
QY	61	AAIAIAGTATAGTASILIFISFTYFTSTWLADCYAPAPVTKGKNYTMQVMSYLGGR	120
Db	52	AAIAIAGTATAGTATVAVPAITTYTSTLADCYSPISITGNTNNTMGVMSYLGGR	111
QY	121	VQLGVAQYQNLIGTAVGTITNLSISLVAVGKSCNCFHDKGHTADCTISNTPYNAFGIIO	180
Db	112	VQLGVAQYVNLVGVITIGYITITASISLVAIIGKSNCYHDKGKACSVSNPYNAAFGIYQ	171
QY	181	VILSQINPFLKLSFLSILMAAVMSFTYATIGIGLATATVAGGKVGKTSMTGTAAGVDVTA	240
Db	172	IILSQINPFLKLSFLSILMAAVMSFTYATIGIGLATATVAGGKVGKTSMTGTAAGVDVTA	231
QY	241	OKIWRSGVAGDIAFAVATVULIEQDTLASSPKNKAKRASLVGVSTTTPFYILIGC	300
Db	232	EKVWKLEQVAGDIAFAVATVULIEQDTLASSPKNKAKRASLVGVSTTTPFYILIGC	291
QY	301	IGYAFGNNAKPGDFLTDGFEFEPFMIIDPANNACIVHLIGAYOVAPIFOFVEKKCNRN	360
Db	292	IGYAFGNNAKPGDFLTDGFEFEPFMIIDPANNACIVHLIGAYOVAPOFQFVENDCNK	351
QY	361	YPDKKFTTSSYSNVNPEFLKRNISLPELWMPRAVYVITTVAMLPFPFNALIGLIGASAF	420
Db	352	WPDKNFTTSSYSNVNPEFLKRNISLPELWMPRAVYVITTVAMLPFPFNALIGLIGASAF	411
QY	421	WPLTVPEPVMHIAOTKIKKYSARMIATKMCYVCLIVSLAAGSIAGLISSVKTYKPF	480

```

Db      412 WPLTFFPVAHMLIAQAKKXKXSRRLMLNLVCLIVLSALAAVGSITGLINSVSKYKF 471
Oy      481 RTM 483
Db      472 KNL 474

RESULT 4
P92934 PRELIMINARY; PRT; 481 AA.
ID      P92934
AC      P92934;
DT      01-MAY-1997 (TREMBlrel. 03, Created)
DT      01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT      01-MAR-2002 (TREMBlrel. 20, Last annotation update)
GN      Antio acid permease 6.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eumids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_Taxid=3702;
[1]
RX      SEQUENCE FROM N.A.
RP      STRAIN=CV. LANDSBERG; TISSUE=LEAF;
RC      MEDLINE=96373203; PubMed=8776904;
RA      Ratsch D., Hirner B., Schmeizler E., Frommer W.B.;
RT      "Salt stress-induced proline transporters and salt stress-repressed
RT      broad specificity amino acid permeases identified by suppression of a
RL      yeast amino acid permease-targeting mutant.";
RL      Plant Cell 8:1437-1446(1996).
[2]
RN      SEQUENCE FROM N.A.
RC      STRAIN=COLUMBIA;
RC      MEDLINE=20181125; PubMed=10718197;
RA      Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA      Tabata S.;
RT      "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT      features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT      clones.";
RL      DNA Res. 7:31-63(2000).
DR      EMBL; X95736; CAA65051.1; -.
DR      EMBL; AB026627; BAA97227.1; -.
DR      InterPro: IPR002422; AA/re1.pmease2.
DR      InterPro: IPR001395; Aldo/ke1_red.
DR      Pfam: PF01490; Ae_trans.1.
DR      PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
SQ      SEQUENCE 481 AA; 53020 MW; DEBICAT348760ADI CRC64;

Query Match 74.3%; Score 1871; DB 10; Length 481;
Best Local Similarity 73.6%; Pred. No. 7,8e-138;
Matches 345; Conservative 60; Mismatches 62; Indels 2; Gaps

Oy      15 ESDGATVVDPTKNDVDEGRKRRTGYMLTNSAHITTAIVGSGVSLAMAIAOLGWTAGTS 74
Db      11 QSPREHIDDTKNDKDEGRKRRTGYMTGSAHITTAIVGSGVSLMAIAOLGWTAGGA 70
Oy      75 ILILFSFIYFTSTMLADCYAPDPVYGRKNTYMDVRSYLGSRKYOLGVAYOYNTLIG 134
Db      71 VLMASFITYFTSTMLADCYSPDPVYGRKNTYMEVRSYLGSRKYOLGGLAOYNTLIG 130
Oy      135 VTYGYTTISISLVAVGASNCFFHDKGHTADCTISNPTPAWVFGIITVYISQITPNFKLSF 194
Db      131 ITTGYYTTISISMVAVKSNCFHNGNINVCATSNTPFITFAIIITLISQITPNFLSM 190
Oy      195 LSLMAAVNSFTATYTGIGLATAIYA-GGKYGKTSMTGTAVGVDVTAQKIMRSFOAVGDI 253
Db      191 LSLIAAVNSFCVASTIGVGLSLAKAAGGEHVRTLTVGYTVGVDSGAEXIMRFQAIQDI 250
Oy      254 AFAYATATVLIETIOTTLASS-PANRKNMKNRASLVGVSTTTPPYTLGCGITGAAPGNNAAG 312
Db      251 AFYAASTVLIETIOTTLAAGPSNKNMKNRASLVGVSTTTPPYTLGCGVGAAGNDADG 310

```


AC G39134.1
PT 01-NOV-1996 (TEMBLrel. 01. Created)
01-JAN-1999 (TEMBLrel. 09. Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20. Last annotation update)
DE Amino acid permease (Amino acid carrier, putative).
GN AAP3 OR P2P24.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
CX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Fischer W.N.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=95332342; PubMed=7608199;
RX Fischer W.N., Kwart M., Hummel S., Frommer W.B.;
RT "Substrate specificity and expression profile of amino acid
RT transporters (AAPs) in Arabidopsis."
RL J. Biol. Chem. 270:16315-16320(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA Fischer W.N.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA STRAIN=CY. COLUMBIA;
RC Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Rongling C.M., Koo H., Fujii C.Y., Dteback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome 1 BAC P2P4 genomic sequence."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: X77499; CAA54630.1; -.
DR EMBL: AC078898; AAC92903.1; -.
DR InterPro: IPR002422; AA/rel_pmaseae2.
DR Pfam: PF01490; Aa_trans; 1.
SQ SEQUENCE 476 AA; 52037 MW; FFC8C42037804F67 CRC64;
Query Match 59.7%; Score 1504; DB 10; Length 476;
Best Local Similarity 57.6%; Pred. No. 3,3e-109;
Matches 273; Conservative 97; Mismatches 100; Indels 4; Gaps 4;
QY 10 NHSTAESGDVYVSDPTKVNDEDGEEKRGTMLTSAHITTVAVTSGVTLSTAAVQLQW 69
DB 4 NHOVLAVD-MPTGSGSKYLDGDKRRKRGSVWTAHITTVAVTSGVTLSTAAVQLQW 62
QY 70 IAGTSILIFSFITYETSTMLADCYRAPDPVTKGNATYVMVSVSYGGRVQLCGVAY 129
DB 63 IAGVWVLLFSATVYFTTSSILAACYRSGDPISGKNTTYMDAVASNLGVKVTLCGIYQ 122
QY 130 GNLIGVAVGTTTASISLVAVGSKNCFHKKHTLACTISNTPYMAVGIQIIVLSQPNR 189
DB 123 LNTFGVAVGTTTASISLVAVGSKNCFHKSQKDCPNNSNPYMTAGLVQIILSQIPDF 182
QY 190 HKLSFSLMAAVNSFYATATIGILAIA-TVAGKGVKGTSMGTAVGVDTVAQKIRWSPQ 248
DB 183 DQLMWLSILAAVNSFYTSAGLALGIAQVAVNGKV-KSLTGISIGA-VTEIGKIWRTPQ 240
QY 249 AVGGIAVAVYAVLVLEIDDTLRSSPAEKMAKRSALVGSVTTTFPYTLCGGCYAAGN 308
DB 241 ALGDIAPVSYSTILLEIQVYKSPSEKTKKATLVSVSYTTFMFWLGCCKGYAIFGD 300
QY 309 NAGPDFLTDFGFEFPFWLIDFANACTAVHLIGAYVFAQPIFOVEKKCKRNRYDKNFT 368
DB 301 LSPGNLITGCGFYNPFWLIDIANAAIVHLIGAYVYCOPIFAPIEKQASIQPDFSEFIA 360
QY 369 SEYSVAVPFGKFNISIFRLVWRTAVVTTTVAAVIPPFAVNIIGLIGASFWFLTYFP 428
DB 361 KDIKIPFGKPLAVNFRIRITRYVITVITVYISLWLFEPDVGGLGAGFWFLTYFP 420
QY 429 VEHIAOTKIKKYSARMIALKTMVCYGLIVSLIAAAGSIAGLISSVTKYRPT 482

[illegible]

```
QY 248 QAVGDIAPAYATVLTIEIOTLRSSPAENKAMKASLAVGVSTTTFFYLGCIGYAAF 307
DB 256 QALGDIAPAYSYSVLIEIOTVRSPPAESKTKWKATKISIAVTTTFYMLCGSMGYAAF 315
QY 308 NNAPGDELTFGFEPEFWLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRNYPDNKFI 367
DB 316 DNAPGNLTGFGFNFWLIDANALVHLVGAYQVFAQPIAFIEKSVAEIPDNDFL 375
QY 368 TSEYSVNVPP-FLGKFNISLRLVWRTAYVVTTVVAMIPIPPFNAILGLIGAAFWPLTVY 426
DB 376 SKEPEIRIPGPKSPYKVNVRVGVSTTTTISMLMPFFNDVVGILGALGFWPLTVY 435
QY 427 PVENHIAOTKIKYASARWALTMCVCLIVSLAAAGSIAGLISSVKTKPRTMH 484
DB 436 PVENYIKQKVERKSTWVCLQMLSVACLIVSVAGVSIAGVWLDLKVYKPKFTY 493

RESULT 9
Q8RWAB PRELIMINARY; PRT; 466 AA.
AC Q8RWAB;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amino acid transporter AAP4.
GN AT5G63850.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Bahr J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY093224; AAM13223.1; -.
SQ SEQUENCE 466 AA; 51359 MW; 11B1D5329C630D89 CRC64;

Query Match 59.4%; Score 1496; DB 10; Length 466;
Best Local Similarity 58.0%; Pred. No. 1.4e-108;
Matches 266; Conservative 99; Mismatches 92; Indels 2; Gaps 2;

QY 27 KNVEDGREGKTGTWLTASAHITAVIGSVLSLAWAIAQLGWIAGTISILLISFTYFT 86
DB 9 KCFDDGRLKRSCTVWTAHAITAVIGSVLSLAWAIAQLGWIAGTISILLISFTYFT 68

QY 87 STMILADCYRAPDPVTGKRNVTYMDVRSYLGGRKVQLCGVAYQGNLIGVTGVTITASIS 146
DB 69 STLLSDCYRTGDPVSGKRNVTYMDAVRSILGGRFKICGLIQYLNLFGITVGTITASIS 128

QY 147 LVAVGKNCNCHDKHTADCTISNYPMAVFGIIQVILSOIPNFHKLFSLSMAAVMSFTY 206
DB 129 MMAIKRSNCNCHESGGKPNCHMSNPYIMFGVTEILLSQIKDFQIWLWSIVAAMSFTY 188

QY 207 ATIGIGLAIATVAGGKVGKTSMTGTAVGVVDVTAQAOKIWRSPQAVGDIAPAFAYATVLEI 266
DB 189 SAIGLALGIIQVAANGVVGKSLTGISGA-VTQTKIWRFTQALGDIAPAFAYSVVLIEI 247

QY 267 QDTRLSSPAENKAMKASLAVGVSTTTFFYLGCIGYAAFGNNAPGDELTFGFEPEFWL 326
DB 248 QDTRSPAESKTKWKATKISIAVTTTFYMLCGSMGYAAFQKAPGNLLTGFGYNPFWL 307

QY 327 IDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRNYPDNKFIITSEYSVNVPP-FLGKFNISL 385
DB 308 LDVANAALVHLVGAYQVFAQPIAFIEKQAAAPFSDSLVTKYEYIRIPGRSPYKVN 367

QY 386 FRLVWRTAYVVTTVVAMIPIPPFNAILGLIGAAFWPLTVYFPVEMHIAQTKIKKYSARW 445
DB 368 FRYRSGFVLTIVTISMLMPFFNDVVGILGALGFWPLTVYFPVEMYIRQKVERMSKW 427

QY 446 IALKTMCVCLIVSLAAAGSIAGLISSVKTKPRTMH 484
DB 428 VCLQMLSCGCLMITLVAGVSIAGVWLDLKVYKPKFTY 466
```

```

RESULT 11
ID 039135 PRELIMINARY; PRT; 466 AA.
AC 039135;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Amino acid transporter.
GN AAP4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid 11; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Fischer W.N.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
MEDLINE=95332342; PubMed=7608199;
RA Fischer W.N., Kwart M., Hummel S., Frommer W.B.;
RT "Substrate specificity and expression profile of amino acid
RT transporters (AAPs) in Arabidopsis."
RL J. Biol. Chem. 270:16315-16320(1995).
DR EMBL: X77500; CAA54631.1;
DR InterPro: IPR002422; AA/rel_pmasee2.
DR Pfam: PF01490; Aa_trans; 1.
SQ
SEQUENCE 466 AA; 51470 MW; 1538DFC32C63B778 CRC64;

Query Match 59.3%; Score 1494; DB 10; Length 466;
Best Local Similarity 58.0%; Pred. No. 1.9e-108;
Matches 266; Conservative 98; Mismatches 93; Indels 2; Gaps 2;

QY 27 KNDVEDREKRTGTMLTASHTITTAIGSCYSLMAIAQLGNTIGTITLISFTFT 86
DB 9 KCFDDRLRSGLTWTASHTITTAIGSCYSLMAIAQLGNTIGTITLISFTFTYS 68
QY 87 STMADCYRAPDVTGKRNVTYMDVRSYLAGRKVQLCGVAQGNLIGVTGYTTASIS 146
DB 69 STLSDCYRGPDPGKRNVTYMDVRSYLAGRKVQLCGVAQGNLIGVTGYTTASIS 128
QY 147 LVAVGKNCNCHDKHTADCTISNTPYMAVFGIIOVILSIPNFKLSLSMAAVSGFTY 206
DB 129 MMAIKRNCNCHESGKKNPCMSNPYMAVFGIIOVILSIPNFKLSLSMAAVSGFTY 188
QY 207 ATIGIGLAIATVAGGKVTSMGTAVGVDTAAQKIMRSPQAVDIAFAVAATVLIET 266
DB 189 SAIGLAIIGIIVANGVKSGLGISIGA-VIOTQKIMRTFQALGDIAFAVSIVLIET 247
QY 267 QDTLRSSPAENKAKRSLVGVSTTFPFTILGCGTGYAFCGNNAFGDPLTFGFEPPWL 326
DB 248 QDVRSPSPASKMKIATRIISAVITTFYMLCGCMGYAFGDKAGNLLTFEGFYNPFWL 307
QY 327 IDFNACIAVHLIGAVQVFOPIFOFEVKKCRNYPDNKFTISEYVNVF-FLGKFNISL 385
DB 308 LDVANAIVIHVGAVGVFOPIFAFIEKQLAARFDDSLVYKEVEIKIIPGRSPYKRV 367
QY 386 FRLYWRTRAYVITTVVAMIFPPFNAILGLIGASFWPLTVFVPEMIAQIKIKKYSARV 445
DB 368 FRAYVRGFAVLTVIVISMLPFNDVVGILGLGFPLTVFVPEMIRQKVERMSWKV 427
QY 446 IALKTCYVCILVSLAAASIGLISVYKYKFFRTM 484
DB 428 VCLQMLSCGCMITLVGVGSINGVWLDKVYKFFRTY 466

```

```

RESULT 12
ID 092PM7 PRELIMINARY; PRT; 379 AA.
AC 092PM7;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)

RESULT 13
ID 092PM7 PRELIMINARY; PRT; 488 AA.
AC 092PM7;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Putative amino acid transporter protein.
GN B1065E10.17.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;

QY 51 AVISGSLSLAMIAQLGNTIGTITLISFTFTSTMLADCTRAPDPVTGKRNITYMD 110
DB 1 AVISGSLSLAMIAQLGNTIGTITLISFTFTSTMLADCTRAPDPVTGKRNITYMD 60
QY 111 VRSYSLGKRVQCGVAQGNLIGVTGYTTASISLVAVGSKNCFDCKHTADCTISNY 170
DB 61 VVKAYLGRKRVQCGVAQGNLIGVTGYTTASISLVAVGSKNCFDCKHTADCTISNY 120
QY 171 PYMAVFGIIOVILSIPNFKLSLS-IAAAMSTYATITIGLAIATVAGGKVKTSMT 229
DB 121 PFMTAFACIOIVSOIPNFKLSLS-IAAAMSTYATITIGLAIATVAGGKVKTSMT 180
QY 230 GTAVGVVTAQKIMRSPQAVDIAFAVAATVLIETDTRSSPAENKAKRSLVGV 289
DB 181 GVAAGIDVSTKRWKTLQIGDIARAVYANVLIETDTRSSPAENKAKRSLVGV 240
QY 290 TTFPFTILGCGTGYAFCGNNAFGDPLTFGFEPPWLIDFNACIAVHLIGAVQVFOPI 349
DB 241 PPHSFYVLCGCMGYAFVYDAPNPLTFGFEPPWLIDFNACIAVHLIGAVQVFOPI 300
QY 350 FOFEKKCRNRYPDNKFITSEYVNVFPGKFNISLFLWRTRAYVITTVVAMIFPPFN 409
DB 301 FAFVGGCRDMPEKNFTIREHIEVRFVGYVYVNLPLVWRITTVIITVAVMLFFFPN 360
QY 410 AILGLIGASFWPLTVFP 428
DB 361 DFLGLIGASFWPLTVFP 379

```

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
clone: B1065E10";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003661; BAB90366.1; -.
SQ SEQUENCE 488 AA; 52871 MW; 853333A9F579FCDB2 CRC64;

Query Match 58.4%; Score 1471.5; DB 10; Length 488;
Best Local Similarity 57.6%; Pred. No. 1.2e-106;

Matches 264; Conservative 93; Mismatches 98; Indels 3; Gaps 3;

QY 26 TKNVDEDEGKRTGTMTASAHITAVIGSGVLSLAWAIAQLGWIAGTSIILIFSFTYF 85

DB 30 SKLFDGDRPRNGTMTASAHITAVIGSGVLSLAWAIAQLGWIAGTSIILIFSFTYF 89

QY 86 TSTMADCYRAPDPVTKGRNNTYMDVRSVLGKVKOLGCVAYGNLIGVTGVTITASI 145

DB 90 TSSLLSDCYRGDVPVTKGRNNTYMDVRSVLGKVKOLGCVAYGNLIGVTGVTITASI 149

QY 146 SLVAVGKSNCFHDKGHTADCTISNYPYMAVFGIIQVILSQIPNFHKLSPLSIMAAMVMSFT 205

DB 150 SWLAIGRANCFHKGHDPCNVSSVPVTVIIVGVAEVPFQIPDFDQISWLSMLAAVMSFT 209

QY 206 YATIGIGLAIAVAGKVGKTSMTGAVGVDTAAOKIWRSSQAVGDIAPAYATVILIE 265

DB 210 YSVIGLSIGVQVWANGKLSGTISIGV-VTPMDKWRSLQAFGDIAPAYATVILIE 269

QY 266 IQDTLRS-SPAENKAMKRASLVGVSTTTFFYLGCIGIYAAGFNAPGDFLTDFGFPFEPF 324

DB 269 IQDTIRAPPSESAMKRAVTVSVAVTVTFYMLCGSMGYAAGFDGAPGNLLTGFGFYEPF 328

QY 325 WLIDFANACIAVHLIGAYOVAQPIPOFVEKKCNRNYPDNKEITSEYSVNV-PFLGKFN 383

DB 329 WLDDIANNAAIVHLIGAYOVAQPIPOFVEKKCNRNYPDNKEITSEYSVNV-PFLGKFN 389

QY 384 SLFRLVWRTAYVITVTVVAMIFPFFNAILGLIGAASFPLTVYFVEMHIAQTIKIKYSA 443

DB 389 NLFRTWRTAFVAVTVTVSMLLPFDFNDVVGFLGALGFPLTVYFVEMHIAQTIKIKYSA 448

QY 444 RWALKMTCVCLVLSLAAAGSTAGLISVTKYKPPR 481

DB 449 RWVCLQMLSVGCLVISIAAAGSTAGVMSDLKVRPFPK 486

RESULT 14

Q93X13

ID Q93X13 PRELIMINARY; PRT; 481 AA.

AC Q93X13.

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Amino acid permease AAP4.

GN AAP4.

OS Vicia faba var. minor (tick bean).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.

OX NCBI_TaxID=3907;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV, FRIBO; TISSUE=SEED COAT;

RA Miranda M.;

RL Thesis (2001).

RL Department of Faculty of Mathematics and Natural Sciences,

RL Martin Luther University of Halle-Wittenberg, Halle (Saale), Germany.

DR EMBL; AJ318811; CAC51425.1; -.

DR InterPro; IPR00422; AA/re1_prmase2.

DR Pfam; PF01490; Aa trans; 1.

SQ SEQUENCE 481 AA; 53420 MW; 46C7A8474C9FC679 CRC64;

Query Match 58.0%; Score 1462; DB 10; Length 481;

Best Local Similarity 55.5%; Pred. No. 6.3e-106;

Matches 269; Conservative 92; Mismatches 110; Indels 14; Gaps 5;

QY 5 NTEGHNHSTAESGDATVSD---PTKNVDEDEGKRTGTMTASAHITAVIGSGVLSLA 61

DB 6 NASNNHQT-----FYVSDIQDQDSKDFDGRVKTGTSTASAHITAVIGSGVLSLA 60

QY 62 WAIQAQLGWIAGTSIILIFSFTYF-TSTMADCYRAPDPVTKGRNNTYMDVRSVLGKVK 121

DB 61 WAIQAQLGWIAGTSIILIFSFTYF-TSTMADCYRAPDPVTKGRNNTYMDVRSVLGKVK 120

QY 122 QLCGVAQYGNLIGVTGVTITASISLVAVGKSNCFHDKGHTADCTISNYPYMAVFGIIQV 181

DB 121 QLCGLIQVNLVGVGIVTVASAIWMSVRSNCFHSGKDPCHMSNNTYMIATFAGVQI 180

QY 182 ILSQIPNFHKLSPLSIMAAMVMSFTVATIGIGLAIAVAGKVGKTSMTGAVGV-DVTAA 240

DB 181 IFSQIPDFDQISWLSMLAAVMSFTYSTIGLGLGKVIENK---KEAGTITGLNDVTKA 236

QY 241 QKWRSSQAVGDIAPAYATVILIEIQDTLRS-SPAENKAMKRASLVGVSTTTFFYLGC 299

DB 237 QKTWGSLSQALGDIAPAYATVILIEIQDTIKAPPSSEKTKKCATLSIVITTFPYMLCG 296

QY 300 CIGYAAFGNNAPGDFLTDFGFPFEPFLIDFANACIAVHLIGAYOVAQPIPOFVEKKCN 359

DB 297 CFGYAAFGNSSPGNLLTGFGFYNPFWLLDIANAALVHLIGAYOVAQPIPOFVEKKCN 356

QY 360 NYDNKEITSEYSVNVVPLGKFNISLFRVWRTAYVITVTVVAMIFPFFNAILGLIGAAS 419

DB 357 RFPDSDFVNDVKVPIPLGDRYKLANFRLVWRTAYVITVTVVAMIFPFFNAILGLIGAAS 416

QY 420 FWPVTVVFPVEMHIAQTIKIKYSAKWIALKTCVCLVLSLAAAGSTAGLISVTKYK 479

DB 417 FWPVTVVFPVEMHIAQTIKIKYSAKWIALKTCVCLVLSLAAAGSTAGLISVTKYK 476

QY 480 PRTMH 484

DB 477 FKTY 481

RESULT 15

Q92RS1 PRELIMINARY; PRT; 466 AA.

ID Q92RS1

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Amino acid carrier.

GN AAP3.

OS Ricinus communis (Castor bean).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.

OX NCBI_TaxID=3988;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SANGUINEAS;

RA Medlam A., Marvier A.C., Hall J.L., Williams L.E.;

RT "Functional characterization and expression analysis of the amino acid

permease RCAAP3 from castor bean."

RL Plant Physiol. 120:1049-1056(1999).

DR EMBL; AJ132228; CAA10608.1; -.

DR InterPro; IPR002422; AA/re1_prmase2.

DR Pfam; PF01490; Aa trans; 1.

SQ SEQUENCE 466 AA; 51003 MW; 7734CD6BFAC53F2F CRC64;

Query Match 57.3%; Score 1442.5; DB 10; Length 466;

Best Local Similarity 57.7%; Pred. No. 2e-104;

Matches 266; Conservative 85; Mismatches 107; Indels 3; Gaps 3;

QY 26 TKNVDEDEGKRTGTMTASAHITAVIGSGVLSLAWAIAQLGWIAGTSIILIFSFTYF 85

DB 7 SKWYDDGKLGKRTGTMTASAHITAVIGSGVLSLAWAIAQLGWIAGTSIILIFSFTYF 66

QY 86 TSTMADCYRAPDPVTKGRNNTYMDVRSVLGKVKOLGCVAYGNLIGVTGVTITASI 145


```
Db 67 TSTILSACVYTPGVNGKXNYTMDAVRNLGGAFFKICGVQVYNLIGVAIGTTIASST 126
QY 146 SLVAVGKSNCFDHPKHTADCTIS-NYPYMAVFGIIOVILSQIPNFHK-LSFSLMAAVMS 203
Db 127 SMAVVRKSNCFHKSFAKNPCMHKCSQSLHDCILLEVNESSQIPDFDQTMGGLSVAAIMS 186
QY 204 FTYATTIGIGLAIVTAVAGKVGKTSMTGTAVGVDTAAQKIMRSFOAVGDIAPAYATVL 263
Db 187 FTYSTIGLIGIAEVTKNGKAMGSMGTISIGT-VTEQKIMRSFOALGDIAPAYSYSLIL 245
QY 264 IETQDITRSSPAPENKAMKQASLVGVSTTFFPILCGCTGYAAGNAPGDFLDFGFFEP 323
Db 246 IETQDITRSSPAPESKTRKATLISVSVTILFYMLCGCEGYAAGFDMSPGNLLTGFGFYNP 305
QY 324 FMLIDFANACIAVHLIGAYOVFAOPIFOFVEKKCNRYPDNKFITSEYSVNVPLGKENI 383
Db 306 YMLDDIANVAIVHLVGAQVYQCPFAFVEKAAVQRPDSEFLIKDIPGCKRYNL 365
QY 384 SLRLVWRITAVVITTVAMIPFENAILGLIGAASFMLTVYFPVEMHIAQTKIKKXSA 443
Db 366 NLFPMWRITVFVITTVISMLLPFENDIVGLGALGFWPLTVYFPVEMYIAQKKIPKXST 425
QY 444 RMIATKTMCVVCLIVSILAAAGSIAGLISSVKTYKPPFRTMH 484
Db 426 RMLCLOITLSAACLIITITAAAAGSIAGVIVDLKTVKPPQTTY 466
```

Search completed: November 22, 2002, 13:31:27
Job time : 37 secs

PS Claim 5; SEQ ID NO 596; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.

XX SQ Sequence 485 AA;

Query Match 100.0%; Score 2519; DB 23; Length 485;
Best Local Similarity 100.0%; Pred. No. 5.9e-266;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSFNTEGHNHSTAESGDATVSDPTKNVDEGDKRTGTWLTASAHITAVIGSVLSL 60

Db 1 MKSFNTEGHNHSTAESGDATVSDPTKNVDEGDKRTGTWLTASAHITAVIGSVLSL 60

Qy 61 AWAIQGLWAGTSSILLIFSFITYFTSTMLADCYRAPDPVTGKKNYTYMDVRSYLGGRK 120

Db 61 AWAIQGLWAGTSSILLIFSFITYFTSTMLADCYRAPDPVTGKKNYTYMDVRSYLGGRK 120

Qy 121 VOLCCVAQYGNLIGVTGVTYTTASISLVAVGKSNCFPHDKGHTADCTISNYPYMAVFGIIQ 180

Db 121 VOLCCVAQYGNLIGVTGVTYTTASISLVAVGKSNCFPHDKGHTADCTISNYPYMAVFGIIQ 180

Qy 181 VILSQIPNPKLSFLSMAAVMSFTYATIGLAIATVAGGKVGKTSMTGTAVGVDTAA 240

Db 181 VILSQIPNPKLSFLSMAAVMSFTYATIGLAIATVAGGKVGKTSMTGTAVGVDTAA 240

Qy 241 QKIWRSPQAVGDI AFAYATVLIIEIQTLSRSPAENKAMKASLVGVSTTTFFYILCGC 300

Db 241 QKIWRSPQAVGDI AFAYATVLIIEIQTLSRSPAENKAMKASLVGVSTTTFFYILCGC 300

Qy 301 IGYAAGFNNAAGDFLTDFGFFPFWLIDFANACIAVHLIGAYQVFAQPIFOFVEKKCNRN 360

Db 301 IGYAAGFNNAAGDFLTDFGFFPFWLIDFANACIAVHLIGAYQVFAQPIFOFVEKKCNRN 360

Qy 361 YPDNKFITSEYSVNVVPLGKFNISLRLVWRTAYVITTVVAMIPFFFNAILGLIGAASF 420

Db 361 YPDNKFITSEYSVNVVPLGKFNISLRLVWRTAYVITTVVAMIPFFFNAILGLIGAASF 420

Qy 421 WPLTVYFPVEMHIAQTKIKKYSARWIALKTCYVCLIVSLAAAGSIAGLISSVKTYKPF 480

Db 421 WPLTVYFPVEMHIAQTKIKKYSARWIALKTCYVCLIVSLAAAGSIAGLISSVKTYKPF 480

Qy 481 RTMHE 485

Db 481 RTMHE 485

RESULT 2

AAR47812

ID AAR47812 standard; Protein; 485 AA.

XX

XX

AC AAR47812;

XX

XX 20-JUL-1994 (first entry)

XX

DE Sequence of amino acid transporter (AAT).

XX

XX Amino acid transporter; Aat; plant transformation.

XX

OS Arabidopsis thaliana.

XX

XX DE4222315-A.

XX

PD 13-JAN-1994.

XX

PF 05-JUL-1992; 92DE-4222315.

XX

PR 05-JUL-1992; 92DE-4222315.

XX

PA (GENB-) INST GENBIOLOGISCHE FORSCHUNG.

XX

PI Frommer W;

XX

XX WPI; 1994-017036/03.

DR N-PSDB; AAQ55058.

XX

XX DNA coding for aminoacid transporter proteins - used for

PT modifying transporter activity in plants

XX

XX Disclosure; Page 12-14; 15pp; German.

XX

CC cDNA from 2-leaf seedlings of A. thaliana was inserted into pFL61
CC and used to transform the Pro-transport deficient yeast strain 22574d
CC Plasmid pFL61-ppp1-20 was isolated from a transformant capable of
CC growing with Pro as the sole N source. AAQ55058 is the coding sequence
CC in pFL61. A notI fragment from this was inserted into p-Bluescript
CC SK to obtain ppp1-20 (DSM 7130).

XX SQ Sequence 485 AA;

Query Match 99.9%; Score 2516; DB 15; Length 485;

Best Local Similarity 99.8%; Pred. No. 1.3e-265;

Matches 484; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSFNTEGHNHSTAESGDATVSDPTKNVDEGDKRTGTWLTASAHITAVIGSVLSL 60

Db 1 MKSFNTEGHNHSTAESGDATVSDPTKNVDEGDKRTGTWLTASAHITAVIGSVLSL 60

Qy 61 AWAIQGLWAGTSSILLIFSFITYFTSTMLADCYRAPDPVTGKKNYTYMDVRSYLGGRK 120

Db 61 AWAIQGLWAGTSSILLIFSFITYFTSTMLADCYRAPDPVTGKKNYTYMDVRSYLGGRK 120

Qy 121 VOLCCVAQYGNLIGVTGVTYTTASISLVAVGKSNCFPHDKGHTADCTISNYPYMAVFGIIQ 180

Db 121 VOLCCVAQYGNLIGVTGVTYTTASISLVAVGKSNCFPHDKGHTADCTISNYPYMAVFGIIQ 180

Qy 181 VILSQIPNPKLSFLSMAAVMSFTYATIGLAIATVAGGKVGKTSMTGTAVGVDTAA 240

Db 181 VILSQIPNPKLSFLSMAAVMSFTYATIGLAIATVAGGKVGKTSMTGTAVGVDTAA 240

Qy 241 QKIWRSPQAVGDI AFAYATVLIIEIQTLSRSPAENKAMKASLVGVSTTTFFYILCGC 300

Db 241 QKIWRSPQAVGDI AFAYATVLIIEIQTLSRSPAENKAMKASLVGVSTTTFFYILCGC 300

Qy 301 IGYAAGFNNAAGDFLTDFGFFPFWLIDFANACIAVHLIGAYQVFAQPIFOFVEKKCNRN 360

Db 301 IGYAAGFNNAAGDFLTDFGFFPFWLIDFANACIAVHLIGAYQVFAQPIFOFVEKKCNRN 360

Qy 361 YPDNKFITSEYSVNVVPLGKFNISLRLVWRTAYVITTVVAMIPFFFNAILGLIGAASF 420

Db 361 YPDNKFITSEYSVNVVPLGKFNISLRLVWRTAYVITTVVAMIPFFFNAILGLIGAASF 420

Qy 421 WPLTVYFPVEMHIAQTKIKKYSARWIALKTCYVCLIVSLAAAGSIAGLISSVKTYKPF 480

Db 421 WPLTVYFPVEMHIAQTKIKKYSARWIALKTCYVCLIVSLAAAGSIAGLISSVKTYKPF 480

Qy 481 RTMHE 485

Db 481 RTMHE 485

RESULT 3

AAG06246

ID AAG06246 standard; Protein; 485 AA.

XX

XX

AC AAG06246;

XX

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 2950.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX
XX 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125768.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130649.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137724.
XX 07-JUN-1999; 99US-0138094.
XX 08-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 10-JUN-1999; 99US-0139119.
XX 14-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 16-JUN-1999; 99US-0139454.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140895.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

```
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 99.0%; Score 2495; DB 21; Length 485;
Best Local Similarity 99.4%; Pred. No. 2.5e-263;
Matches 482; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKNTECHNHSTAESGDATVSDPTKNVDEGKREKRTGTLTASAHIIITAVIGSVLSL 60
DB 1 MKNTECHNHSTAESGDATVSDPTKNVDEGKREKRTGTLTASAHIIITAVIGSVLSL 60

QY 61 AWAIAQLGWIAGTSILLIFSFITYFTSTMLADCYRAPDPVTKGRNYYTMDVVRSYLGGK 120
DB 61 AWAIAQLGWIAGTSILLIFSFITYFTSTMLADCYRAPDPVTKGRNYYTMDVVRSYLGGK 120

QY 121 VOLCGVAQXNLTGVTVGTGVTITASISLVAVGKSNCPHDKGHTADCTISNYPYMAVFGIIQ 180
DB 121 VOLCGVAQXNLTGVTVGTGVTITASISLVAVGKSNCPHDKGHTADCTISNYPYMAVFGIIQ 180

QY 181 VILSQIPNFHKLSPLSIMAAVMSTYATIGIGLAIAATVAGGKVGKTSMTGTAVGVDVTA 240
DB 181 VILSQIPNFHKLSPLSIMAAVMSTYATIGIGLAIAATVAGGKVGKTSMTGTAVGVDVTA 240

QY 241 QKWRSFQAVDITAFAYATVLTIEQDTLRSSPAENKAMKEASLVGVSTTTTFFYILGCG 300
DB 241 QKWRSFQAVDITAFAYATVLTIEQDTLRSSPAENKAMKEASLVGVSTTTTFFYILGCG 300
```

```
QY 301 IGVAAFGNNAFGDFTDGFPPFWLIDPANACIAVHLIGAYQVFAQPIQFQVEKKNRN 360
DB 301 IGVAAFGNNAFGDFTDGFPPFWLIDPANACIAVHLIGAYQVFAQPIQFQVEKKNRN 360
QY 361 YPNKFTITSEYSVNVFPLGKFNISLFRLVWRTAYVVTVMIFPPFNAILGLIGAASF 420
DB 361 YPNKFTITSEYSVNVFPLGKFNISLFRLVWRTAYVVTVMIFPPFNAILGLIGAASF 420
QY 421 WPLTVPPVPMHIAQTKIKKYGARWIALTKWCYVCLIVSLAAAGSIAGLISSVKTYKPF 480
DB 421 WPLTVPPVPMHIAQTKIKKYGARWIALTKWCYVCLIVSLAAAGSIAGLISSVKTYKPF 480
QY 481 RTMHE 485
DB 481 RTMHE 485

RESULT 4
AAG06247
ID AAG06247 standard; Protein; 397 AA.
AC AAG06247;
XX
DI 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2951.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
```

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140053.
PR 23-JUN-1999; 99US-0140054.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141827.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142190.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146387.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149823.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152163.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154179.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.

```

PR 29-OCT-1999; 99US-0162142.
Query Match 81.4%; Score 2056; DB 21; Length 397;
Best Local Similarity 99.7%; Pred. No. 1.7e-215;
Matches 396; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 89 MLADCYRAPDPVTCKRNTYMDVVRSLGKRVQLCGVAQYGNLIGVTGTTTASISLV 148
DB 1 MLADCYRAPDPVTCKRNTYMDVVRSLGKRVQLCGVAQYGNLIGVTGTTTASISLV 60
OY 149 AVGSNCNCFHDKGHTADCTISNYPYMAVFGIIQVILSQIPNFHKLFLSFLSMAAVMSFTYAT 208
DB 61 AVGSNCNCFHDKGHTADCTISNYPYMAVFGIIQVILSQIPNFHKLFLSFLSMAAVMSFTYAT 120
OY 209 IGIGLATATVAGGKVGKTSMTGTAAGVGVDTAAQKWRSPQAVGDIAPAYATVLIIEQD 268
DB 121 IGIGLATATVAGGKVGKTSMTGTAAGVGVDTAAQKWRSPQAVGDIAPAYATVLIIEQD 180
OY 269 TLRSPPAENKAMKRAASLVGVSTTTFFVILGCGICGYAAFGNNAAGDPLTDFGFPFPLWID 328
DB 181 TLRSPPAENKAMKRAASLVGVSTTTFFVILGCGICGYAAFGNNAAGDPLTDFGFPFPLWID 240
OY 329 PANACIAVHLICAGVQFAQPIFOFVEKKCNRPDKNPFITSEYSVNVPLGKFNISLFL 388
DB 241 PANACIAVHLICAGVQFAQPIFOFVEKKCNRPDKNPFITSEYSVNVPLGKFNISLFL 300
OY 389 VWRtayVVITTVVAMIPFFFNAILGLIGAAFPWPLTVYPPVEMHIAQTKIKKYSARMIAL 448
DB 301 VWRtayVVITTVVAMIPFFFNAILGLIGAAFPWPLTVYPPVEMHIAQTKIKKYSARMIAL 360
OY 449 KTMCVVCLIVSLLAAGSIAGLISSVKYKPPRTMHE 485
DB 361 KTMCVVCLIVSLLAAGSIAGLISSVKYKPPRTMHE 397

RESULT 5
ABB90933
ID ABB90933 standard; Protein; 475 AA.
XX
AC ABB90933;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 144.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PP 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 144; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
XX (ABB90790-ABB94016) for herbicidally active compounds, comprising
XX aligning and comparing nucleic acid or amino acid sequences from plant
CC
```

```

CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 475 AA;
Query Match 78.4%; Score 1974.5; DB 23; Length 475;
Best Local Similarity 75.4%; Pred. No. 1.7e-206;
Matches 364; Conservative 58; Mismatches 52; Indels 9; Gaps 2;
OY 1 MKNSTEGHNSHTAESGDVATVSDPTKNVDEGDKRKRTGTLTASAHITIAVIGSVLSL 60
DB 1 MDATN----NPSAVESGDA----AVKSVDGDKRKRTGTLTASAHITIAVIGSVLSL 51
OY 61 AWATAQLGWIAGTSIILIFSYFTSTMLADCYRAPDPVTCKRNTYMDVVRSLGGRK 120
DB 52 AWATAQLGWIAGTSIILIFSYFTSTMLADCYRAPDPVTCKRNTYMDVVRSLGGRK 111
OY 121 VOLCGVAQYGNLIGVTGTTTASISLVAGKSNCFHDKGHTADCTISNYPYMAVFGIIQ 180
DB 112 VOLCGVAQYGNLIGVTGTTTASISLVAGKSNCFHDKGHTADCTISNYPYMAVFGIIQ 171
OY 181 VILSQIPNFHKLFLSFLSMAAVMSFTYATIGLAIATVAGGKVGKTSMTGTAAGVDTAA 240
DB 172 IILSQIPNFHKLFLSFLSMAAVMSFTYATIGLAIATVAGGKVGKTSMTGTAAGVDTAA 231
OY 241 OKIWRSPQAVGDIAPAYATVLIIEQDTLRSSPAENKAMKRAASLVGVSTTTFFVILGCG 300
DB 232 EKWKLFQALGDIAPAYATVLIIEQDTLRSSPAENKAMKRAASLVGVSTTTFFVILGCG 291
OY 301 IGYAAGFNNAAGDPLTDFGFPFPLWIDPANACIAVHLIGAYQVFAQPIFOFVEKKCNRN 360
DB 292 IGYAAGFNNAAGDPLTDFGFPFPLWIDPANACIAVHLIGAYQVFAQPIFOFVEKKCNRN 351
OY 361 YPDNKEITSEYSVNVPLGKFNISLFLVWRtayVVITTVVAMIPFFFNAILGLIGAAFP 420
DB 352 WQSNFINKESYSSKVPPLGKCRVNLFLVWRtayVVITTVVAMIPFFFNAILGLIGAAFP 411
OY 421 WPLTVYPPVEMHIAQTKIKKYSARMIALKTMCVVCLIVSLLAAGSIAGLISSVKYKPP 480
DB 412 WPLTVYPPVEMHIAQTKIKKYSARMIALKTMCVVCLIVSLLAAGSIAGLISSVKYKPP 471
OY 481 RTM 483
DB 472 KNL 474

RESULT 6
AAG06248
ID AAG06248 standard; Protein; 377 AA.
XX
AC AAG06248;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2952.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PP 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
```


PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 16-MAY-1999; 99US-0134370.
PR 16-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145214.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148345.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151108.
PR 31-AUG-1999; 99US-0151133.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152633.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 77.1%; Score 1942; DB 21; Length 377;
 Best Local Similarity 99.7%; Pred. No. 4.4e-203;
 Matches 376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 109 MDVRSYLGKVKQLCGVAGYGNLIGVTGYTTTASISILVAVGKSNCFHDKGHTADCTIS 169
 DB 1 MDVRSYLGKVKQLCGVAGYGNLIGVTGYTTTASISILVAVGKSNCFHDKGHTADCTIS 60
 QY 169 NYPMAVFGIIQVLSQIPNFHKLSPISLMAAVMSPTYATIGIGLAIATVAGGKVGKTSM 228
 DB 61 NYPMAVFGIIQVLSQIPNFHKLSPISLMAAVMSPTYATIGIGLAIATVAGGKVGKTSM 120
 QY 229 TGTAVGVDTAAQKINRSFOAGVDIAPAYATVLTIEIOTLRSSPAENKAMKASLVGV 208
 DB 121 TGTAVGVDTAAQKINRSFOAGVDIAPAYATVLTIEIOTLRSSPAENKAMKASLVGV 180
 QY 289 STTFFVILCGICGYAAGFNAPGDELTDGFPPEPWLIDFANACIAVHLIGAYQVFAQP 348
 DB 181 STTFFVILCGICGYAAGFNAPGDELTDGFPPEPWLIDFANACIAVHLIGAYQVFAQP 240
 QY 349 IFQVEKKCNRYPDNKFITISEYVNVVPELGFKNISLFLVWRTAYVITTVVAMIPEFF 408
 DB 241 IFQVEKKCNRYPDNKFITISEYVNVVPELGFKNISLFLVWRTAYVITTVVAMIPEFF 300
 QY 409 NAILGLIGAAASFNLTYVFFVEVHIAQTIKKYSARWIAKTMCVCLIVSLAAAGSIA 468
 DB 301 NAILGLIGAAASFNLTYVFFVEVHIAQTIKKYSARWIAKTMCVCLIVSLAAAGSIA 360
 QY 469 GLISSVTKYKPFRTMHE 485
 DB 361 GLISSVTKYKPFRTMHE 377

RESULT 7
 ABB93755
 ID ABB93755 standard; Protein; 481 AA.
 XX
 AC ABB93755;
 XX

DT 31-MAY-2002 (first entry)
 XX Herbicidally active polypeptide SEQ ID NO 2966.
 XX Herbicidal; plant; agriculture; herbicide.
 XX Arabidopsis thaliana.
 XX WO200210210-A2.
 XX 07-FEB-2002.
 XX 28-AUG-2001; 2001WO-EP09892.
 XX 28-AUG-2001; 2001WO-EP09892.
 XX (FARB) BAYER AG.
 XX Tietjen K, Weidler M;
 XX WPI; 2002-269010/31.
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 PS Claim 5; SEQ ID NO 2966; 261pp + Sequence Listing; English.
 XX The invention relates to identifying target proteins
 CC (ABB9790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 XX Sequence 481 AA;
 SQ
 Query Match 74.3%; Score 1871; DB 23; Length 481;
 Best Local Similarity 73.6%; Pred. No. 3.6e-195;
 Matches 345; Conservative 60; Mismatches 62; Indels 2; Gaps 2;
 QY 15 ESGDAYTVSDPTKNVEDGDKRKTGTWLTASAHITAVIGSVLSLAWAIAQLGWIAGTS 74
 DB 11 QSPPEHIGDTNKNFDEGDRKRTGTWLTGSAHITAVIGSVLSLAWAIAQLGWIAGPA 70
 QY 75 ILLIFSPITYFTSTMLADCYRAPDPVTGKRNVTYMDVRSYLGKVKQLCGVAGYGNLIG 134
 DB 71 VLMAFSPITYFTSTMLADCYRSPDPVTGKRNVTYMEVRSYLGKVKQLCGVAGYGNLIG 130
 QY 135 VTGCTTTASISILVAVGKSNCFHDKGHTADCTISNYPMAVFGIIQVLSQIPNFHKLSP 194
 DB 131 ITIGYTTTASISILVAVGKSNCFHKGHNKCATSNTPPMIIFALIIQLSQIPNFHLSW 190
 QY 195 LSIMAAVMSPTYATIGIGLAIATVA-GKVGKTSMTGTAVGVDTAAQKINRSFOAGVDI 253
 DB 191 LSILAAVMSFCYASIGVLSIAKAAAGGGEHVRTTLTGTVDIGIVSGAEKIWRTPQAI 250
 QY 254 AFAYATVLTIEIOTLRSSPAENKAMKASLVGVSTTFFVILCGICGYAAGFNAPG 312
 DB 251 AFAYATVLTIEIOTLKGAPSEKAMKASLVGVSTTFFVILCGICGYAAGFNAPG 310
 QY 313 DFLTDFGFFPEPWLIDFANACIAVHLIGAYQVFAQIPQFVEKKCNRYPDNKFITSEYS 372
 DB 311 NFLTGFGEFFPEPWLIDFANACIAVHLIGAYQVFCQIPQFVESQSAKRWPDNKFITGEYK 370
 QY 373 VNVPELGFKNISLFLVWRTAYVITTVVAMIPEFNAILGLIGAAASFNLTYVFFVEMH 432
 DB 371 IHVPECCDGFSEINFLRLVWRTSYVVTAVVAMIFPEFNDFLGLIGAAASFNLTYVFFVEMH 430

OY 433 IAQTKIKKYSARWIAATKMCYCLIVSLAAGSINAGLISVYKTKPR 481
Db 431 IAQKIKKFSFTWTKILSWTCFIVSLVAAAGSVQGLIQSLKDKPRQ 479

RESULT 8
AAG28804
ID AAG28804 standard; Protein; 476 AA.
XX AAG28804;
AC
XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 34160.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX EPI033405-A2.
PV
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PR
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 26-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	23-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161982.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity 59.7%; Score 1505; DB 21; Length 476;		
Matches 273; Conservative 97; Mismatches 100; Indels 4; Gaps		
Qy	10	NHSTAESDAYTVSDPTKNVDEGREKRTGTMLTASAHITAVICSGVLSLAWATAQIGW 69
Db	4	NHGTUAVD-MQTCGSKYLDGGRKNTSVWFTASAHITAVICSGVLSLAWATAQIGW 62
Qy	70	IAGTSILLIFSPITYFTSTMLADCYRAPDPVTGKNTYMDVVRSLGGRKVQLCGVAQY 129
Db	63	LACPVWMLFVSVVFTSSLLAACVRSQDPISGKRNVTYMDAVRSLGGRKVQLCGVQY 122

	Qy	130	GNLIGVTGYGTYTASISLVAVGKSNCFDHDGTADCTTSINPYNAVFGLIIQVILSQIPNF	189
	Dd	127	: : : : : : : : : : : : : : : : : :	
	Dd	123	LNIFGVAIGYTIASAIENMAIKRSNCFHKSGKGOPCHMNSNPYMIAFGLVLIQSIPDF	182
	Qy	190	HKLISFLSMAAAMSFYTATIGIGLAIA-TVAGKGVKGTSMTGTAVGVDVTTAAOKIWRSFQ	248
	Dd	183	: : : : : : : : : : : : : : : : : :	
	Dd	183	DQLWLSILAAMSFITYSSAGLAGIAQQVVNGKV-KGSLGISIGA-VTE TOKIWRTQ	240
	Qy	249	ANGDTAFAYAYATVLIEIOPTLRSSPAENKAKRASLIVGUSTTFPFVLLCGCIGYAAFGN	308
	Dd	241	: : : : : : : : : : : : : : : : : :	
	Dd	241	ALGDFAFYASYIILEIOPTVKSPSEKTKWKATLVSVSTTMFMVLCGCMGYAAF GD	300
	Qy	309	NAPGDFLD FGFEPFPFWLLIDFANACIAVHLIGAYQVFAQPIQFQVEKKCNRNYPDNKFIT	368
	Dd	301	: : : : : : : : : : : : : : : : : :	
	Dd	301	LSPGNLLAGFGYNPYPWLLDIANAALVIHLIGAYQVYCPLFAFIEQAASI QFPDSEFTA	360
	Qy	369	SYSVNVPFLCKFNISLFRLVWTATVTVTTTTVAMI PFFFNAILGLIGAASFPLTVTP	428
	Dd	361	: : : : : : : : : : : : : : : : : :	
	Dd	361	KDKIPIEGFKPLKLANFRILTMTVEVII TTVISM LPPFDNDVVLLGALGFPLTVTP	420
	Qy	429	VEMHIAQT KIKKY SARMI ALKTWCYYCLIVSLLAAAGSIAGLISSVKTYKPFT	482
	Dd	421	: : : : : : : : : : : : : : : : : :	
	Dd	421	WE MYIAQKK IPRWSTRWVCIQVFSLGCLVVSIAAAAAGSIAGVLDDLKSYKPPRS	474
	RESULT 9			
	ABBS91631			
	ID	ABB91631 standard; Protein; 476 AA.		
	XC	ABB91631;		
	XT			
	DT	31-MAY-2002 (first entry)		
	DE	Herbicidally active polypeptide SEQ ID NO 842.		
	Xx	Herbicidal; plant; agriculture; herbicide.		
	Kw	Arabidopsis thaliana.		
	OS			
	Pn	WO200210210-A2.		
	Pd			
	Pf	07-FEB-2002.		
	Ff	28-AUG-2001; 2001WO-EPO9892.		
	Pf			
	Pr	28-AUG-2001; 2001WO-EPO9892.		
	Pa	(FARB) BAYER AG.		
	Pi	Tietjen K, Weidler M;		
	Xx			
	Dd	WPI; 2002-269010/31.		
	Pt			
	Pt	Identifying plant target proteins for herbicidally active compounds,		
	Pt	comprising aligning and comparing nucleic acid or amino acid sequences		
	Pt	from plant with nucleic acid or amino acid sequences from non-plant		
	Pt	organisms -		
	Xx			
	Ps	Claim 5; SEQ ID NO 842; 261pp + Sequence Listing; English.		
	Xc			
	Cc	The invention relates to identifying target proteins		
	Cc	(ABBS90790-ABS94016) for herbicidally active compounds, comprising		
	Cc	aligning and comparing nucleic acid or amino acid sequences from plants		
	Cc	with nucleic acid or amino acid sequences from non-plant organisms us-		
	Cc	suitable search parameters, where plant sequences having an E-value		
	Cc	greater by a factor of 3 than the E-value of most similar non-plant		
	Cc	sequences are selected. The polypeptides or nucleic acids encoding the		
	Cc	are useful for identifying modulators. The identified modulators are		
	Xx			
	Sq	Sequence 476 AA;		

XX	Sequence	476 AA;
5Q		

Query Match 59.7%; Score 1504; DB 23; Length 476;
Best Local Similarity 57.6%; Pred. No. 4,4e-155;
Matches 273; Conservative 97; Mismatches 100; Indels 4; Gaps 4;

OY 10 NHSTAEAGDAYTSDPTKQVNDGKERKGTGTMILASAHITITAVIGSGVLSIAAMAIAQLGW 69
DB 4 NHQTVLAVD-MPQGGSKYLDGDKNKRKGTGWTASAHITITAVIGSGVLSIAAMAIAQLGW 62
OY 70 IAGTILIFSEFTYFTSTMLADCYRAPDPVTKKNTYTMQVRSYLGGRVVOUGVAOY 129
DB 63 IAGVVMVLLTSATLYFTSSILACYSRGPISGRKNTYTMQVRSYLGGRVVOUGVAOY 122
OY 130 GNLIGTVGTTTASISLVAVGKSNCFPHDKHTADCTISNYPMAVPSIIIVILISQIPNF 189
DB 123 LNIQVVAIGYTIASISMAIKRSNCFHSGKDPCHMNSNYPMAVPSIIIVILISQIPNF 182
OY 190 HKLSFSLTMAVNSFTYATIGIGLATA-TVAGSGKSTMTGTAVGVNTAAKIMPSFO 248
DB 183 DQMMVLSITLAAVNSFTYSSAGLIGIAQVAVNGKV-KSSLTISIGLAVTETQKIMRTFO 240
OY 249 AVSDIAFAVAVTAVLRIQDTLSSPAENKAMKRAVSGVSTTFEPYILGCTIGYAAFGN 308
DB 241 ALGDIAPAVSYSTILIEIDTVKSPSEKTKMKAIVSVSTIMFVMLCGMGYAAFGD 300
OY 309 NAGPDLTDFGFEFPEFLIDFANACIAVHLIGAYVPAQPIQFQVEKKCNHYPDNKFT 368
DB 301 LSPGNLTGEGFNPYMLDIANAALVHILGAYQVYCPLEAFIEKQASIQFDPSEPIA 360
OY 369 SEYVNVNPFLEKFNISLFLWRTAVAVITTVAMIFPEFNAIIGLIGAASFMPLYTFP 428
DB 361 KDIPITIPGEPFLINFLIRMTVEIITTVISMILPEFNVVGLLALGMPLYTFP 420
OY 429 VEMHIAOTKIKKYSARWIALKTCQVCLIVSLAAAGSIAGLISVTKYKPEPT 482
DB 421 VEMTIAOKKIFRMSTRVNCIQVSLGLVSIAAAGSIAGVLLDKYKPEPS 474

RESULT 10
AAG28805
ID AAG28805 standard; Protein; 464 AA.
XX AAG28805;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 34161.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hydriatisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0101439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134570.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139459.
PR 17-JUN-1999; 99US-0139492.
PR 17-JUN-1999; 99US-0139493.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.

XX 07-FEB-2002.
PD 28-AUG-2001; 2001WO-EP09892.
XX 28-AUG-2001; 2001WO-EP09892.
PR 28-AUG-2001; 2001WO-EP09892.
XX (FARB) BAYER AG.
XX Tietjen K, Weidler M;
XX WPI; 2002-269010/31.
XX
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
XX Claim 5; SEQ ID NO 2615; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
XX Sequence 493 AA;
SQ
Query Match 59.6%; Score 1501; DB 23; Length 493;
Best Local Similarity 56.7%; Pred. No. 9,8e-155;
Matches 271; Conservative 94; Mismatches 111; Indels 2; Gaps 2;
QY 8 GHNHSTAESGDATVSDPTKNVDEGDERKRTGTWLTASAHITTAIVIGSGVLSLAWAIAQL 67
Db 17 GHQVFDVASHDFPQPAFCFDDGRLKRTGTWMTASAHITTAIVIGSGVLSLAWAIAQL 76
QY 68 GWIGTSLILIFSTITFSTMTADCYRAPDPVTGKENTVDVRSYLGKRXQLGVA 127
Db 77 GWIGTSLILIFSTITFSTMTADCYRAPDPVTGKENTVDVRSYLGKRXQLGVA 136
QY 128 QYGNLIGVTGYTTTASISLVAVGKSNCFHDKGHTADCTISNRYMAVFGIIOVILSGIP 187
Db 137 QYNLFGIAGYTTAASISWMAIKRSNCFHSGKDPCHMSNPMTIVFGVABILLSGVP 196
QY 188 NFHLSTSLSTMAAMSTFTATIGLAIATVAGKYGKTSMTGTVAGVDVTAQKIRSF 247
Db 197 DFDIMWISIVAAWSTYSATIGLAIQVAVANGFKSLGISIG- VTDQTKIMTIF 255
QY 248 QAVGDIAPAYAVATVLIETDITRSPBAENKMKRASLVGVSSTTFYILGCGIYAFG 307
Db 256 QALDIAFAVSYSVLIEIDTVRSPBAESKTKATKISIVATITFIMLCGSMGYAFG 315
QY 308 NNAAGDELTCGFEPEFMLIDFANACIAVHLIGAHOVEOPFQVEKKCRNPDNKEI 367
Db 316 DAAGNLTGEGFNPLMDIANAAIVHLVGAIVOPAPQIFATIEKSVAKRIPDNDEL 375
QY 368 TSEYSVNVP-FLGKFNISLFRVWKRTAYVITTVAMIRPFENALIGIGAASEMPLTVY 426
Db 376 SKEBEIRIPGKSPYKVNVRMYRSGFVTTTISMLPFENDVIGLIGALGEMPLTVY 435
QY 427 FPMVMTIAQCTIKKYASRMATLKMTCVCLIVSLAAASITGLISVYTKYKPPRTM 484
Db 436 FPEVMTIKQKVEKSTRWVCLQMLSVACLIVSVAGVSIAGVMDLVKTKKPFASY 493

AC AAG42351;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 52809.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0128845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135153.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.

PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	24-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	28-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	03-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	08-JUL-1999;	99US-0142803.	PR	28-SEP-1999;	99US-0156458.
PR	08-JUL-1999;	99US-0142920.	PR	29-SEP-1999;	99US-0156596.
PR	12-JUL-1999;	99US-0142977.	PR	04-OCT-1999;	99US-0157117.
PR	13-JUL-1999;	99US-0143542.	PR	05-OCT-1999;	99US-0157753.
PR	14-JUL-1999;	99US-0143624.	PR	06-OCT-1999;	99US-0157865.
PR	15-JUL-1999;	99US-0144005.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	16-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158293.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159284.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159330.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159638.
PR	20-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0144814.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160767.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160778.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160815.
PR	22-JUL-1999;	99US-0145192.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160989.
PR	23-JUL-1999;	99US-0145224.	PR	25-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	26-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.			
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR					
PR	</				


```

Oy 267 ODTLRSSPAENKAMKRASLVGSTTFFYLGGCIGYAAFGNNAAGDPLDTGFPPEPWL 326
Db 248 ODTVRSPAESKTKMTKATRIISIAVTTTFYMLGCMGYAFAFGKAGNLLTGFGYNNPWL 307
Oy 327 IDPANACIAVLHIGAYOVFAOPIFQFVEKKCNRNYPDNKFTISEYVNP-FLGKFNISL 385
Db 308 LDVANAALVHILVGAAYOVFAOPIFAFIEKQAAARFPDSDLTYKEIEIRIPGFRSPYKYNV 367
Oy 386 FRLVWRDVAVVITTVVAMIPFPFNAILGLIGAASFWPLTVYFPEVMHIAQTKIKKYSARW 445
Db 368 FRAVYRSGFVVLTVVISMLMPFNDVVGILGALGFWPLTVYFPEVMYIRORKVERMSKMW 427
Oy 446 TALKTCVCYCLIVSLAAGSIAGLISVYKTYKPFRTMH 484
Db 428 VCLQMLSCGLMITLVAGVSIAGVMDLKVYKPFRTTY 466

RESULT 13
ABB93937
ID ABB93937 standard; Protein; 466 AA.
AC ABB93937;
XX 31-MAY-2002 (first entry)
DT 31-MAY-2002 (first entry)
XX 31-MAY-2002 (first entry)
DE Herbicidally active polypeptide SEQ ID NO 3148.
KM Herbicidally active polypeptide SEQ ID NO 3148.
XX Herbicidal; plant; agriculture; herbicide.
KM Herbicidal; plant; agriculture; herbicide.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN MO200210210-A2.
XX 07-FEB-2002.
PD 07-FEB-2002.
XX 28-AUG-2001; 2001MO-EP09892.
PF 28-AUG-2001; 2001MO-EP09892.
XX 28-AUG-2001; 2001MO-EP09892.
PR 28-AUG-2001; 2001MO-EP09892.
XX (FAR) BAYER AG.
PA (FAR) BAYER AG.
XX Tietjen K, Weidner M;
PI Tietjen K, Weidner M;
XX WPI; 2002-269010/31.
DR WPI; 2002-269010/31.
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX Claim 5; SEQ ID NO 3148; 261pp + Sequence Listing; English.
PS Claim 5; SEQ ID NO 3148; 261pp + Sequence Listing; English.
XX The invention relates to identifying target proteins
XX (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX Sequence 466 AA;
SQ Sequence 466 AA;

Query Match 59.3%; Score 1495; DB 23; Length 466;
Best local similarity 58.0%; Pred. No. 4,1e-154;
Matches 266; Conservative 98; Mismatches 93; Indels 2; Gaps 2;

```

```

Db 69 STLSDCYRCDPVSIGKRNVTYMDVRSILGGFRFKCLLOYLNFQITVGYTTAASIS 128
Oy 147 LVAVGKSNCFHDKGHTADCTISNYPMAVFGIIQVILSOIPNFHKLSTLSMAAVMSTTY 206
Db 129 MMAIKRSNCPIHESGKKNPCIMSSNPYIMRFGYTEILLSQIKDPDQIMWLSIVAAIMSTTY 188
Oy 207 ATIGGLAIIATVAGKAGKXTSMGTAVGVDTAAOKIRASFQAVGDIFAVAAYATVLEI 266
Db 189 SAIGLALGIIQVANGVKSILGISIGA-VTQTKIKRTFPALGDIAFAISYSVLEI 247
Oy 267 ODTLRSSPAENKAMKRASLVGSTTFFYLGGCIGYAAFGNNAAGDPLDTGFPPEPWL 326
Db 248 ODTVRSPAESKTKMTKATRIISIAVTTTFYMLGCMGYAFAFGKAGNLLTGFGYNNPWL 307
Oy 327 IDPANACIAVLHIGAYOVFAOPIFQFVEKKCNRNYPDNKFTISEYVNP-FLGKFNISL 385
Db 308 LDVANAALVHILVGAAYOVFAOPIFAFIEKQAAARFPDSDLTYKEIEIRIPGFRSPYKYNV 367
Oy 386 FRLVWRDVAVVITTVVAMIPFPFNAILGLIGAASFWPLTVYFPEVMHIAQTKIKKYSARW 445
Db 368 FRAVYRSGFVVLTVVISMLMPFNDVVGILGALGFWPLTVYFPEVMYIRORKVERMSKMW 427
Oy 446 TALKTCVCYCLIVSLAAGSIAGLISVYKTYKPFRTMH 484
Db 428 VCLQMLSCGLMITLVAGVSIAGVMDLKVYKPFRTTY 466

RESULT 14
AAG13704
ID AAG13704 standard; Protein; 493 AA.
AC AAG13704;
XX 17-OCT-2000 (first entry)
DT 17-OCT-2000 (first entry)
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13297.
KM Arabidopsis thaliana protein fragment SEQ ID NO: 13297.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX EP1033405-A2.
XX 06-SEP-2000.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.

```

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 59.2%; Score 1492; DB 21; Length 493;
 Best Local Similarity 56.5%; Pred. No. 9.5e-154;
 Matches 270; Conservative 94; Mismatches 112; Indels 2; Gaps 2;

QY 8 GHNHSTAESGDATVSDPTKNDVEDGERKGTGWTLSAHIIITVIGSGVLSLAMAIAQL 67
 DB 17 GHQVFDVASHDPPPOPAKCPDDGRLKRTGTWTASAHIIITVIGSGVLSLAMAIAQL 76
 QY 68 GWIAGTSLILIFSTFTYTSMLADCYRAPDPVTKRNYTYMDVVRSYLGGKXVQLGVA 127
 DB 77 GWIAGPAMVMLFSLVTLVLSSTLSDCYRTGDAVSGKRYTYMDAVRSILGGFKFKICGLI 136
 QY 128 QYGNLIGVTGYTITASISLVAVGKSNCFHGDHATDCTINPYMAVFGIIQVILSOIP 167
 DB 137 QYLNLFYIAIGYITLAIASISMLAKSNCFHKGSGKDPCHMSNPMYVFGVABILLSQVP 196
 QY 188 NFHKLSFLSMAAVMSFTYATIGIGLAIATVAGKVGKTSMTGAVGVDTAAQKIRSF 247
 DB 197 DFOIWMISIVAAVMSFTYSAIGLAIQVAVANGVFKSGSLTGISIGT-VTQOKIRTF 255
 QY 248 QAVGDIAFAAVATVLEIOTLRSSPAENKMRASLVGVSTTFPFYILGCGIGYAFG 307
 DB 256 QALGDIAFAISYVLEIOTLRSPASKTKMKAATKISIAVTTFPMLGSGMVAFG 315
 QY 308 NNAPEGDLTDFGFEFPMWIDPANACIAVHLIGAYQVFAQPIFOFEKKCRNYPDNKFI 367
 DB 316 DAABGNLITGFGYVPMWIDIANALIVHLVGAQVFAQPIFAVEIKSVARYPDNDFL 375
 QY 368 TSESVAVV-FLCKENISLFLVWTRAVVITTVVAMPPFNALIGLIGASFPWLTYY 426
 DB 376 SKFEINIDPFKSPYKRVKFRVYRSGFVAVTTVISMMPFNDVVGILGALGFWPLTVY 435
 QY 427 FPEVEMHIAQTKIKKYSARWIALKTMCYCLIVSLAAAGSIAGLISVYKYPRTWH 484
 DB 436 FPEVEMYIKORKEKMSRWYCLQMLTVACIVISVAVGSIAGVWMLDKVYKPKSTY 493

RESULT 15
 ABB91267
 ID ABB91267 standard; Protein; 476 AA.
 AC ABB91267;
 DT 31-MAY-2002 (first entry)
 DB Herbicidally active polypeptide SEQ ID NO 478.
 KW Herbicidal; plant; agriculture; herbicide.
 OS Arabidopsis thaliana.
 PN WO200210210-A2.
 PD 07-FEB-2002.
 PF 28-AUG-2001; 2001WO-EP09892.
 PR 28-AUG-2001; 2001WO-EP09892.
 PA (FARB) BAYER AG.

PI Tietjen K, Weidler M;
 XX WPI; 2002-269010/31.
 DR
 XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX
 PS Claim 5; SEQ ID NO 478; 261np + Sequence listing; English.
 CC
 CC The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 CC
 SQ Sequence 476 AA;

Query Match 55.8%; Score 1406.5; DB 23; Length 476;
 Best Local Similarity 55.8%; Pred. No. 2e-144;
 Matches 256; Conservative 91; Mismatches 101; Indels 11; Gaps 4;

QY 30 DEDGERKGTGWTLSAHIIITVIGSGVLSLAMAIAQLGWIAGTSLILIFSTFTYTSML 89
 DB 21 DDDRPKRTGTWTASAHIIITVIGSGVLSLAMAIAQLGWIAGPAMVMLFSLVFPYSTL 80
 QY 90 LADCYRAPDPVTKRNYTYMDVVRSYLGGKXVQLGVTGYTITASISLVA 149
 DB 81 LCSCRGSDVTKRNYTYMDAHSNLGSIKVGCVQVYVNLPTALGTYITLAIASLV- 139
 QY 150 VGSNCFHGDHATDCTINPYMAVFGIIQVILSOIPNFHKLSFLSMAAVMSFTYATY 209
 DB 140 --TSCQMMGPNDPCHVNGVNYMLAFGIQVIFSOIPDFOIWMISIVAAVMSFAYSAL 196
 QY 210 GIGLAIATVAGKVGKTSMTGAVG-----DVTAAQKIRSFQAVGDIAPVAVATVLI 264
 DB 197 GUGSGVSKVENKEIKGSLTGVTGVTLSGTVTSQIKIWRTFQSLAIATVAYSMTLI 256
 QY 265 EIOTLRSSPAENKMRASLVGVSTTFPFYILGCGIGYAAFNNAPGDFLTDFGFEFP 324
 DB 257 EIOTLVKSPBAEVNTRKATFVSAVTTFVYMLCCCVGAAFGDNAPGNLLAHGCFRNPY 316
 QY 325 WLIDPANACIAVHLIGAYQVFAQPIFOFEKKCRNYPDNKFTISEVAVPFIQK-FNI 383
 DB 317 WLIDIANALIVHLVGAQVQVCEPLPAFVEKESRRPSEFVTKIQL-FQKRPNL 375
 QY 384 SLFLVWTRAVVITTVVAMPPFNALIGLIGASFPWLTYYPEVEMHIAQTKIKKXSA 443
 DB 376 NLFLVWTRFVWTTLISMLMPFNDVVGILGALGFWPLTVYFPEVEMYIAQKVPWNGT 435
 QY 444 RWIALKTMCYCLIVSLAAAGSIAGLISVYKYPRT 482
 DB 436 KWVCLQVLSVTCLPVSVAAAGSVIGVSDLVKVPFQS 474

Search completed: November 22, 2002, 13:29:57
 Job time : 41 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2002, 13:30:02 (Search time 12 Seconds
(without alignments)
632.985 Million cell updates/sec

Title: US-09-854-562-2

Perfect score: 2519 1 MKSFTEGHNSTASGDAY.....SIAGLSVYTKYKFRTHHE 485

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100480 seqs, 15661496 residues

Total number of hits satisfying chosen parameters: 100480

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/BCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	388	15.4	468	10	US-09-860-232A-8
2	212.5	8.4	88	10	US-09-860-232A-18
3	150.5	6.0	504	10	US-09-919-497-67
4	143.5	5.7	547	10	US-09-789-919-66
5	130.5	5.2	547	10	US-09-776-705-4
6	128	5.1	525	10	US-09-871-503-2
7	128	5.1	525	10	US-09-940-919-2
8	124.5	4.9	547	10	US-09-776-705-2
9	122.5	4.9	547	10	US-09-860-232A-5
10	112.5	4.5	506	10	US-09-776-705-5
11	111.5	4.4	619	10	US-09-741-153-2
12	107	4.2	440	10	US-09-815-242-13429
13	104	4.1	1212	9	US-09-981-351-16
14	101.5	4.0	551	9	US-09-895-913A-268
15	96	3.8	2516	10	US-09-817-514A-2
16	95.5	3.8	336	10	US-09-815-242-13409
17	95	3.8	411	10	US-09-815-242-1128
18	95	3.8	462	10	US-09-815-242-10161
19	94.5	3.8	284	10	US-09-800-103-8

20	94.5	3.8	284	10	US-09-800-103-20	Sequence 20, Appl
21	94.5	3.8	310	10	US-09-800-103-30	Sequence 30, Appl
22	94.5	3.8	481	10	US-09-815-242-10979	Sequence 10979, A
23	93	3.7	542	10	US-09-801-368-48	Sequence 48, Appl
24	92.5	3.7	680	10	US-09-996-194-10	Sequence 10, Appl
25	91	3.6	441	10	US-09-950-071-2	Sequence 2, Appl
26	91	3.6	482	10	US-09-815-242-12941	Sequence 12941, A
27	91	3.6	619	10	US-09-741-153-4	Sequence 4, Appl
28	89.5	3.6	290	10	US-09-800-103-10	Sequence 10, Appl
29	89.5	3.6	290	10	US-09-800-103-22	Sequence 22, Appl
30	89.5	3.6	463	10	US-09-815-242-10144	Sequence 10144, A
31	89.5	3.6	619	10	US-09-800-065-2	Sequence 2, Appl
32	89	3.5	392	10	US-09-800-103-4	Sequence 4, Appl
33	89	3.5	392	10	US-09-800-103-16	Sequence 16, Appl
34	89	3.5	418	10	US-09-800-103-26	Sequence 26, Appl
35	89	3.5	430	10	US-09-800-103-12	Sequence 12, Appl
36	89	3.5	430	10	US-09-800-103-24	Sequence 24, Appl
37	89	3.5	456	10	US-09-800-103-34	Sequence 34, Appl
38	88.5	3.5	477	10	US-09-815-242-13840	Sequence 13840, A
39	88	3.5	466	10	US-09-815-242-13831	Sequence 13831, A
40	87.5	3.5	247	10	US-09-800-103-32	Sequence 32, Appl
41	87.5	3.5	289	10	US-09-886-468-24	Sequence 24, Appl
42	87.5	3.5	438	10	US-09-815-242-5129	Sequence 5129, Ap
43	87.5	3.5	658	10	US-09-815-242-11824	Sequence 11824, A
44	87	3.5	484	10	US-09-815-242-12272	Sequence 12272, A
45	86	3.4	439	10	US-09-815-242-5660	Sequence 5660, Ap

ALIGNMENTS

RESULT 1
US-09-860-232A-8
Sequence 8, Application US/09860232A
Patent No. US2002028494A1
GENERAL INFORMATION:
APPLICANT: Curtiss, Roy A. J.
TITLE OF INVENTION: 57256 AND 58289, NOVEL HUMAN
TITLE OF INVENTION: TRANSPORTERS AND USES THEREOF
FILE REFERENCE: 381552001500
CURRENT APPLICATION NUMBER: US/09/860,232A
CURRENT FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: US 60/205,288
PRIOR FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 468
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Consensus amino acid sequence
US-09-860-232A-8
Query Match 15.4%; Score 388; DB 10; Length 468;
Best Local Similarity 30.4%; Pred. No. 2.5e-28;
Matches 139; Conservative 67; Mismatches 183; Indels 68; Gaps 20;
QY 66 QLGMIAGTSLILFFSPT--YFTSTMLADCYRAP-DPYTGKRN--YTYMDVVR-SYLG 118
DB 17 QLGMPICGVLLVLLVGFITLYGHTGTLTKVCEEBEYVPCREBAKSYLLGOSANG 76
QY 119 RKVQTCGV---AQYGLIGVTGYTTSISLVAGVS-NCF-----HDKGHAD--C 165
DB 77 KGLILTFPVGTVGQVYVNFVGNIGYIILAGDLLPKTISLFFCLFLIYQGVNAGVSGD 136
QY 166 -TISNYPMAVF-GIIIVILSIQIPNFHKS-----FLSTMAVMSFTY-ATIGIG 212
DB 137 DLISGNSWIIIFAILITTPSLFIPAPNLLSASNLSTLVLISSISAFSSLAISISFL 196
QY 213 LAITVAGS-----GKVGKISMTGTAAGVDVTAQKIMSFOAVGDIARAVAYATVLEIQ 267
DB 197 IVAVAVIGIFVLGAVYKILMSVETLATVPSYTKLTGLFALIGITVFAFGHAYLPIQ 256


```

Oy 105 -----NYTMDVVRSLG-----117
Db 146 FGMWGRKIGARISITMONIGAMSSYLFITIKELPEVIRAPMGLIEBNTGWIANGVILFV 205
Oy 118 --GRKVOICGVAOYGNLIGVYGYTTASISLVAAGKSNCF-----HDKGTADC 165
Db 206 SVGIILPLSLKMLGYL-GYTSQFSLSCWVFEVSIVYKKFQIPCPALDHNNG---NL 261
Oy 166 TISNYPMAVFGIIIOVILSQIPNFHKLSTLMAAVMSFTYATIGLAIATVAGKVK 225
Db 262 TERN-----LPIHMSISLPDSSSGVNFM---MDYAHN-PAGIDEQVAG-----304
Oy 226 TSMGTAVGVDTAAQKIW-----RSFOAVGDIAPAVAYATVLEIQTLLRSPAEN 277
Db 305 -PLHNSGVEVEAOGAEKCPKPFVFNSTRVAIPIIAFAVCHPEVLEIYSELND--RSR 361
Oy 278 KAMKRALVGVSTTFEYIICGIGYAAFGNNAPODPLDGFEPFMLIDFANACIAVH 337
Db 362 RKQOTVSNISISGMLWYLLAALFGYLSFYGEDELLHAYS-----KYTFDTALLMR 416
Oy 338 LIGAYOV-FAOPIFOVEKKCNENYDNKFTSEYVAVVP--LGFENISLFLWRTAY 394
Db 417 LAVLAATLVLPVILF-----PIRFSVITLLFPKKPFESWLGHFGIA-----AII 460
Oy 395 VVITTVAMIFPFNAIILGIGASFPWLTVPVPMHIAQTKIKKYSARWIAKTMCV 454
Db 461 IALNNIIVILVPIKTYIFGFIASATMLIPLPAFYL--KLWKEPPLSPOKIGALV 517
Oy 455 CLIVSLIAAGSIAGLI 471
Db 518 FLVLTGILFMGSMALII 534

```

RESULT 5

```

US-09-776-705-4
; Sequence 4, Application US/09776705
; Patent No. US2002008219A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001010
; CURRENT APPLICATION NUMBER: US/09/776,705
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-776-705-4

```

Query Match 5.2%; Score 130.5; DB 10; Length 547;

Best Local Similarity 17.9%; Pred. No. 0.00024;

Matches 92; Conservative 87; Mismatches 215; Indels 119; Gaps 16;

```

Oy 32 DGRKRGTWLTSAHITLVIGSSVLSLAAIAQLGMAAGSIIILISFTTSTMLA 91
Db 68 DEHHPGTISGMSFNLISALWGSILGLSYAMANTGIVLFEVIMLLTVAIISLVSALL 127
Oy 92 DCYRAPDPVTGR-----NYTMDVVRSLG-----117
Db 128 KTAKEGSLIYEKLGKAPGMPKIGAFISITMONIGAMSSYLFITIKELPEVIRVEMGL 187
Oy 118 -----GRKVOICGVAOYGNLIGVYGYTTASISLVAAGKSNCF- 156
Db 188 EBNTEGWIANGVILFVSVGIILPLSLKMLGYL-GYTSQFSLSCWVFEVSIVYKKFQ 246
Oy 157 -----HDKGTADCITSNYPMAVFGIIIOVILSQIPNFHKLSTLMAAVMSFTYA 207
Db 247 IPCPLVILDHNNG---NLTENNITLPMHVI-----MLPNNSESTGNFPM---VDYTHR 292

```

```

Oy 208 TIGTGLAIATVAGKVKGKTSMTGTAVGVDTAAQKIW-----RSFOAVGDIAPAVAY 259
Db 293 D-PEGLDEKRAAG-----PLHSGVEVEAHSQDKCPKRYFVFNSTRVAIPIIAFAV 345
Oy 260 ATVLEIQTLLRSPAENKAMKRALVGVSTTFEYIICGIGYAAFGNNAPODPLDFG 319
Db 346 HPEVLPIYSELND--RSRRMQVSNISITGMLWYLLAALFGYLSFYGEDELLHAYS 403
Oy 320 FFEFPMLIDFANACIAVHIGAYOV-FAOPIFOVEKKCNENYDNKFTSEYVAVVP 378
Db 404 -----KYTFDTALLMRKALVAVATLVLPVILF-----PIRFSVITLLRPPFS 449
Oy 379 GKFNISLFLWRTAVVITTVAMIFPFNAIILGIGASFPWLTVPVPMHIAQTKI 438
Db 450 WKHFGIAIIT-----IALNNIIVILVPIKTYIFGFIASATMLIPLPAFYL---KL 501
Oy 439 KKYSAWIAKTMCVCLIVSLIAAGSIAGLI 471
Db 502 VKKEPPLSPOKIGALVPLVLTGILFMGSMALII 534

```

RESULT 6

```

US-09-871-503-2
; Sequence 2, Application US/09871503
; Patent No. US20020076758A1
; GENERAL INFORMATION:
; APPLICANT: George Christien Terstappen
; APPLICANT: Cinzia Felicitia Sala
; TITLE OF INVENTION: POLYPEPTIDE
; FILE REFERENCE: OG 1013
; CURRENT APPLICATION NUMBER: US/09/871,503
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: GB 0013239.9
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-871-503-2

```

Query Match 5.1%; Score 128; DB 10; Length 525;

Best Local Similarity 19.8%; Pred. No. 0.0004;

Matches 100; Conservative 79; Mismatches 228; Indels 98; Gaps 20;

```

Oy 7 EGHNSTASGDAVYVSDPTKNV---DEGGRKGTWLTSAHITVAVIGSVLSIAW 62
Db 80 EGDHIGRSGGAPLPBSGSKDVGSGGFERGHDKRTAMERAGNVVTAIQMVLGLPY 139
Oy 63 AIAQGMIACTIILIFSFITYFTSTMLADCY-RAPDPVTGKNRYTYMDVRS-----114
Db 140 AILHGYL-GLFILIIPAAVVCCYTGKILTACLYEENDEGEVVRVDSVLAAMACAPRF 198
Oy 115 -YLGRVQICGVAOYGNLIGVYGYTTASISLVAAGKSNCFHDKGTADCITSNPYM 173
Db 139 PTLGRV---NVAQIIEVMTCILYVV-----SGULMNSPGL 236
Oy 174 AV---FGII-OVILSQIPNFHKLSTLMAAVMSF---TYATIGTGLAIATVAGKV 223
Db 237 FVSOKWSIIATAVLL-----PCAFILNKLAVSKFSLCTLAHFVINIIVICYLSRA 289
Oy 224 GKTSMTGTAVGVDTAAQKIWRSFO-AVGDIAPAVAYATVLEIQTLLRSPAENKMKR 282
Db 290 RDNAMKVKFYIDV-----KKFPISIGIYFTSTQIFLPSLEGNQ-QPSEFHCMMN 341
Oy 283 ASLVGSTTFEYIICGIGYA---AFGNNAPODPLDGFEPFMLIDFANACIAVH 337
Db 342 WTHIAACVLKGLPALVAAYLTWDETKEVITDNLPGSIRAVNVIF---LVAKALLSYPLP 397
Oy 338 LIGAYOVFAOPIFOVEKKCNENYDNKFTSEYVAVVPFLGKFNISLFLWRTAVVVI 397

```

Db 398 FFAAVEVLEKSLFO-----EGGRAPPACYS-----GDGRKSWGLTLRCALVVF 442
QY 398 TTVVAMIFFPFFNAILGLIGAASFPLTVVFPVEMHIAQTOKIKKYSARM-----IALKTM 451
Db 443 TLLMAIVVPHFALLMGLTGLTGAFLCLLPSLFLH---RLLRKLLMHQVFFDVAIFVI 499
QY 452 CYVCLIVSLLAAAGSIAGLISSVKT 476
Db 500 GGIC---SVSGFVHSLEGLIAYRT 521

RESULT 7

US-09-940-919-2
; Sequence 2, Application US/09940919
; Patent No. US20020082390A1
; GENERAL INFORMATION:
; APPLICANT: Frieddie, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: Same
; FILE REFERENCE: Same
; CURRENT APPLICATION NUMBER: US 09/940,919
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/230,178
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 525
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-940-919-2

Query Match 5.1%; Score 128; DB 10; Length 525;
Best Local Similarity 19.8%; Pred. No. 0.0004;
Matches 100; Conservative 79; Mismatches 228; Indels 98; Gaps 20;

QY 7 EGNHSTAESDAYTVSDPTKNV-----DEGREKRTGTWLTASAHIIITAVIGSGVLSLAW 62
Db 80 EGDHYQRGSGAPLPPEGSKDQVGGGEGFGGHDKPIITAEAGWNTWAIQMFVLGUPY 139
QY 63 AIAQLGIAGTSSILLIFSFITYFTSTMLADCY--RAPDPVTGKRTNYTMDVVR-- 114
Db 140 AILHGGYL-GLFLIIFAVVCYTGKILIACTLYEENEDGEVVRVDSYVAIANACCAPRF 198
QY 115 -YLGRKVOLCGVAQYGNLIGVTGVTITASISLVAVGKSNCFHDKGHTACTTISNYPYM 173
Db 199 PTLGRVV--NVAQIITELVWTCILYVV-----SGNLMNNSPGL 236
QY 174 AV----FGIIL--QVILSQIPNFHKLGFSLMAAVMSF----TYATIGIGLAIATVAGKV 223
Db 237 PVSQKSNLSIATAVLL-----PCAPFLKNLKAVSFSLCTLAHFVINILVIAYCLSR 289
QY 224 KTSMTGTAVGVDTAAQKIRSFQ-AVGDIAPAYAYATVLEIODTLRSSPAENKAMKR 282
Db 290 RDNAWEKVKFYIDV-----KKPISIGLIVPSYTSQIFLESNGNQ-QPSEFHCMMN 341
QY 283 ASLVGVSTTFFYLCCGICGYA-----AFGNNAAGDFLTDGFFPFWLIDFANACIAVH 337
Db 342 WTHIAACVLKGLFALVAYLTWADETKEVITDNLPGSIRAVNVIF----LVAKALLSYPLP 397
QY 338 LIGAYQVPAQPIFOFVEKKCNRYPNKFTITSEYVNVVFLGKFNISLFRLVWRTAYVI 397
Db 398 FFAAVEVLEKSLFO-----EGGRAPPACYS-----GDGRKSWGLTLRCALVVF 442
QY 398 TTVVAMIFFPFFNAILGLIGAASFPLTVVFPVEMHIAQTOKIKKYSARM-----IALKTM 451
Db 443 TLLMAIVVPHFALLMGLTGLTGAFLCLLPSLFLH---RLLRKLLMHQVFFDVAIFVI 499
QY 452 CYVCLIVSLLAAAGSIAGLISSVKT 476
Db 500 GGIC---SVSGFVHSLEGLIAYRT 521

RESULT 8

US-09-776-705-2
; Sequence 2, Application US/09776705
; Patent No. US20020082191A1
; GENERAL INFORMATION:
; APPLICANT: CUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001010
; CURRENT APPLICATION NUMBER: US/09/776,705
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Human
US-09-776-705-2

Query Match 4.9%; Score 124.5; DB 10; Length 547;
Best Local Similarity 17.7%; Pred. No. 0.00088;
Matches 98; Conservative 90; Mismatches 191; Indels 174; Gaps 20;

QY 32 DGREKRTGTWLTASAHIIITAVIGSGVLSLAWAIAQLGMIAGTSSILLIPSFITYFTSTMLA 91
Db 68 DEHPGTTSPGMSFNLSNAINMGSGILGLSYAMAYTGVLFIIMLLAVAILSLYSVHLL 127
QY 92 DCYRAPDPVTGKR-----NYTYMDVVRSYLG- 117
Db 128 KTAKEGSLIYEKLGKAFGNPKGTGAFVITMQNIGAMSSVLPFIKYLEDEVIRAFMGL 187
QY 118 -----GRKVQLCGVAQYGNLIGVTGVTITASISLVAVGKSNCFH 157
Db 188 BENTGYNLNGNVLIIIPVSVGIIPLSLKNLGYL-GVTSFGSLTCMVFVSV----VY 242
QY 158 DKGH-----TADCTISNYPYMAVFGIIQVILSQIPNFHKLGFSLSMAAVMSFTYATIGI 211
Db 243 KKFQPCPLPVLDHSVGNLSFNNTLPMHVVM-----PNNSSESDVNF-----MDYTHRN-PA 295
QY 212 GLAIATVAGKVQKTSMTGTAVGVDTVTAQKI-----WRSFQAVGDIAPAYAYATVL 263
Db 296 GL-----DENQAKGSLHDSGVEYEAHSDCKEPKYFVFNSTAYATPILVFAFVCHPEV 349
QY 264 LEIQDTLRSSPAENKAMKRAKSLVGYSTTTTFFYLCCGICGYAAGFNNAAGDFLTDGFFEP 323
Db 350 LPYISEKD--RSRRKQMTVSNISITGMLVMYLLAALFGY-----LTFYGEVD 396
QY 324 FWLIDFANACIAVHLIGAYQVPAQPIFOFVEKKCNRYPNKFTITSEYVNVVPLGKFN 383
Db 397 -----ELLHAY-----SKVYTLIDIPLL-MVRL 417
QY 384 SLFLVWRTAVVI-----TTVWAMIFP-----PFNA-----I 411
Db 418 AVLVAVTQTPVLPVLPFIRTSVITLFPKPPPSWIRHFLIAVLIANNVLVLVPTIKYI 477
QY 412 LGLIGAASFPLTVVFPVEMHIAQTOKIKKYSARWIALTCMCYCLIVSLLAAAGSIAGLI 471
Db 478 PGFICASSATMLIFILPAVFLYKLVKETFRS---PQKVGALILFLVVGIFPFMIGSMA-LI 533
QY 472 SSVKTYKPFRTMH 484
Db 534 IIDWIYDPPNSKH 546

RESULT 9

US-09-860-232A-5
; Sequence 5, Application US/09860232A
; Patent No. US20020028494A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.


```

; TITLE OF INVENTION: 57256 AND 58289, NOVEL HUMAN
; TITLE OF INVENTION: TRANSPORTERS AND USES THEREOF
; FILE REFERENCE: 381552001500
; CURRENT APPLICATION NUMBER: US/09/860,232A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/205,288
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-232A-5

Query Match      4.9%; Score 122.5; DB 10; Length 547;
Best Local Similarity 18.0%; Pred. No. 0.0014;
Matches 94; Conservative 90; Mismatches 226; Indels 112; Gaps 18;

OY 32 DGRKRTGTLTSAHITTVVIGSVLSLMAIAQLGWTAGTSLILFSTFTTSTMLA 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 DEHPRTTSGMSFNLNIMSGILGLSYAMANTGILLFILMLAAVAILSLYSVHLT 127

OY 92 DCYAPDPVTGKR-----NTYMDVVSYLQ- 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 KTAEGGSLIYEKLGKAPGMPKIGAFVSTIMONGAMSSYLFIKLEEVIRAFMGL 187

OY 118 -----GRVQLCGVAQVGNLIGVTGTTTASISLVAQKSNCFH 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 EENTGEWYLNGLIIFVSVGILLPLSLKMLGYL-GYTSFGSLCWMFVSV-----V1Y 242

OY 158 DKGH-----TADCTISNRYMAVFGIIQVILSQIPNFHKLSPFLSMAVMSFTYATIGI 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 KKFQICPPLPDLHSVGNLSFNNTLPMHYVNL-----PNNSSESDVAFM--WDYHNRN-PA 295

OY 212 GLAATATAGGKVGKTSMTGTAAGVDYTAQKI-----WSPQAVGDIAFANAYATVL 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 GL-----DENQAGSLHDSGVEYEAHSDKCEPKYFVNSRTAYAIPIVFAFAPCHPEV 349

OY 264 IEIDTRSSPAEKAMKASLVGVSTTFPFYILGCGICGYAFAFGNNAAGDEFLDPGPFEP 323
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 LPIYSELKD--RSRRKQVYNSISITGMVMTLALRGLYLFGEVDELL--HAYSK 404

OY 324 FWLIDFANACIAVHLIGAYOVFAOPIFQVEKKCNKNYPDNKFTISEXVNVPLGKFN 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 405 VYTDIDPLMVRALVAV--TLTPVPLVLPPIR-----TSVITLLFP--KRP 448

OY 384 SLPR-LVWRATYVITTVAMIPFPFNAILGLIGASFWPLTVVPEVMTAQTIKKYS 442
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 449 SWHHPLAATLILNNVAVLVPTIKYIFGFIGASSATMLFILPAVFIKLKVKETFR 508

OY 443 ARWIALKTMVCVCLIVSLAAGSIAGLISVYKVPFTTH 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 509 S---POKYGALIFLVGIFPMISMA-LIIDIWIDYDPNSKH 546

RESULT 10
US-09-776-705-5
; Sequence 5, Application US/09776705
; Patent No. US20020082191A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO01010
; CURRENT APPLICATION NUMBER: US/09/776,705
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 506
; TYPE: PRT

```

```

; ORGANISM: Human
US-09-776-705-5

Query Match      4.5%; Score 112.5; DB 10; Length 506;
Best Local Similarity 18.8%; Pred. No. 0.01;
Matches 99; Conservative 75; Mismatches 203; Indels 151; Gaps 20;

OY 2 KSFPTGHNSTAESGDATVSDPTKNVDEDEREKRTGTLTASAHITTVAVIGSVSLA 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 KKYETEPHPTTSGMSFNLN-----ALVSGGILGLS 93

OY 62 WAIQLQGIAGTSLILF--SFITFTSTMLADCYRAPDPVTGKRNITVMVNSYIGR 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 94 YAMANTG-IALFIILLTFVSIFFSLYSVHLTLTANEGG-----SLVEQLQGYK 140

OY 120 KYQLCG-VAQVGNL-----IGVTGTYTTASISLVAV-----GKSNCFHDKHTADCT 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 AFGVLKTLAASGISTMQNIGAMSSYLFIKVELPLVIGALTNIEKTKGLAWLNG----- 194

OY 167 ISNRYMAVFGIIQVILSQIPNFHKLSPFLSMAVMSFTYATIGI-----GLAATV 218
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 --NYLVILVSLVYLPLSLFRNGYLGYSGLLQWFELLVYCKKQVPCVEBALI 252

OY 219 AGKVGKTSMTGT-----AVGVDTA-----AQKIRSFQAVGDIAFANAYATV 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 253 INERTINTTLQPTALVPALSHNTENDSCRPHYFIENGQTVY-----AVPILIFSVCHPA 308

OY 263 LIETQDTRSSPAEKAMKASLVGVSTTFPF-----YLCGCTGYAFAFGNNAAGDEFLTD 317
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 309 VLPFYEEL-----KDRSRRMNVSKISFFAMELMTLALRGLYLFGEVHSESLHT 361

OY 318 FGPF--EPPWLDIFANACIAVHLIGAYOVFAOPIFQVEKKCNKNYPDNKFTISEXSVN 374
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 362 YSILGTDILLIYRLAVLMAVTLVPIVIF-----PIRSSVTHLCAS 405

OY 375 VPFLGKFNISLFLWRTAVVI-----TTVYAMIPFPFNAILGLIGASFWPLTVVPEV 429
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 406 KDFW-----WHSILTVSILAFNLVLVIPPVTRIDFGEFIGAASMLFIILPS 455

OY 430 EMHIAQTK-----IKKYSARWIALKTMVCVCLIVSLAAGSIRGLI 471
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 456 APYIKLVKKEPKMSVQKIGLFR-----LLSGVLMTGSMALLIV 494

RESULT 11
US-09-741-153-2
; Sequence 2, Application US/09741153
; Patent No. US20020102637A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO01059
; CURRENT APPLICATION NUMBER: US/09/741,153
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Human
US-09-741-153-2

Query Match      4.4%; Score 111.5; DB 10; Length 619;
Best Local Similarity 20.5%; Pred. No. 0.017; Indels 143; Gaps 22;
Matches 98; Conservative 64; Mismatches 174;

OY 13 TAESGDATVSDPTKNVDEDEREKRTGTLTASAHITTVAVIGSVLSLMAIAQLGWTAG 72
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 TLESQMET-----RLAFGLSTLDLVALGVGSLTLAGVYVLAQVADK--AG 63

OY 73 TSILIFSFITFTSTMLADCYR--APDPVTGKRN-YTYMDV-----VRSY- 115

```

Db 64 PSIVICF-LVAALSVLAGLCYAFGARPCSGSAYLSYVTVGBLNAFTTGMNLIISYV 122
QY 116 LGGRKVLQCGVAQYQVYVTTTASISLVAVGKSCFHDKGHTADCTISNYPYMAV 175
Db 123 IGTASVARAWSSAFDNLIGNHISKTLOGSIAL-----HVPH--- 158
QY 176 FGIQVILSOIPNFKLSFISIMAWNSP-----TYATIGIGLATATVAGGKVKTS 227
Db 159 -----VLAEPDEFALGLVLLTGLLAGASALVTKVTFGVNLLV-----LGFVM 205
QY 228 MTGTAVGVDDVTAQKIMRSFQAVGDIATAFAYAVATVLIBIQDILRSSPAENKAMKRASLVG 287
Db 206 ISGFVKG-DV-----HNNKLTEDYELAMA-----ELNDTYSGLPLSGGCFVPGFEG 252
QY 288 V-STTTFPYLLOCIGYAFAGNNAQDGLTDFGPEPFWLIDPANAC-IAVHLIGAYQV 344
Db 253 ILRGAATCFYAFVFGDCIATTBGEAQNORS-----IPMGIVISUSVCFLAYFAVSSALT 307
QY 345 PAQIFQVVERKCNRYPDNK-----FITSEYSNVNVPFLGKFNIS----- 384
Db 308 LMPYVQL-----QPESPLPEAFLYIGWAPARYVVAVAGSLCALSTSLGSMFMPPRV 359
QY 395 -----LRLVWR-----TAVVITTVVAMIFPPFNAIL-----GLIGAASFPLTVY 426
Db 360 IYMAEDGLLFRVLARIHTGTRPTIATVVGSIIAAFMAFLPKLTDLVDLMSIGTLAY 418

RESULT 12

US-09-815-242-13429
; Sequence 13429, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13429
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13429

Query Match 4.2%; Score 107; DB 10; Length 440;
Best Local Similarity 20.3%; Pred. No. 0.028;
Matches 88; Conservative 69; Mismatches 174; Indels 102; Gaps 21;

QY 49 ITAVIGSG-VLSLAWATAQGLG-----WIAGTSILLIPSFTYFTSTMLADCYRAPDP- 99

Db 67 LASTVGTGNTIIGVATATKVGPGALFMMWMAA-----FPGMATKYAEGLLAKYRTKDDH 121
QY 100 ---VTGKRNTYMDVVRSY-----LGGRKVLQCGVAQYCNLIGVT--VGYYTITASIS 146
Db 122 GAVAGGPMHYILLGMGEKWRPLADLPVAVAGVLVALLGIGTFTQVNSITESQNTTITSPA 181
QY 147 LVAVGKSCNCFDGHADCTISNYPYMAVFG-----IIOVILSOIPNFKLSFISIMAVM 202
Db 182 ITAL-----VLTTFVAIAVFGGLKSIKSVSTTVVP-----FMAIIIVILG 220
QY 203 SFTYATIGIGLATATVAGG-----KVGKTSMTGTAVGVDDVTAQKIMRSFQAVGDI 253
Db 221 TLTVPFNICKIPGTIALGFTSAFSLAAVGGFAGASVRMAIQNGVARGVPSNESGLGSA 280
QY 254 AFAYATVATVLIBIQDILRSSPAENKAMKRASLVGVSTTTPFYILCGCIGY-----AAGFN 308
Db 281 PIAAAAAK-----NEPVEQGLI---SMTGTFIDTL--IICTLTGLTILVTGWSG 326
QY 309 NAPGDFLTDFGFEPPFWLIDFANACIAVHLIGAYQVFAQPIFQFVERKCNRYPDNKFIT 368
Db 327 DLNGVALTQSAFSTVF--SHFGPALLTIFLV-----LFAFTTILGMWRYGRCF-- 373
QY 369 SEYSNVNVPFLGKFNISLFRVWRTAVVITTVVAMIFPPFNAILGLIGAASFPLTVYFP 428
Db 374 -EPLFGVRFIWLRYVFLVLMVLLGGFIEL--NMVWIIADIVNAIMAL---PNLIALLVMSP 428
QY 429 VEMHIAQTKIKKY 441
Db 429 VV--IAET--KKY 437

RESULT 13

US-09-981-353-16
; Sequence 16, Application US/09981353
; Patent No. US20020160382A1

GENERAL INFORMATION:

; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 1212
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1285632CD1
US-09-981-353-16

Query Match 4.1%; Score 104; DB 9; Length 1212;
Best Local Similarity 17.5%; Pred. No. 0.2; Mismatches 233; Indels 194; Gaps 27;
Matches 109; Conservative 87

QY 6 TEGHN-----HSTAESGD-----AYTVSDPTKNVDEGDKRKTGTWLTASA 46
Db 212 TFGHTMDAVPRIDHVRHTAAQLGKLLRPSLAELHDELEKEPEFDGFGANGEESTPTRDA 271
QY 47 HIITAVIGSGVLSLAW-----AIAQLGWIA-----GTSILLIPSFTYF 85
Db 272 VVTYAESKGVVRFGWIKGVLRCLMNTWGMFLFRLSVIVGQAGIGLVLVIMMATVVT 331
QY 86 TSTMLADCYRAPDP-VTGKRNVTYMDVVRSYLGGKRVQCGVAQYGNLIGVTVGYTTAS 144
Db 332 TITGLSTSAIATNGFVRGGGAY-----YLISRLG---PEFGAIGLIFAFANAVA 379
QY 145 ISLVAVGKSCNCFHD--KGTADCTISNYPYMAVFGIIOVILSOIPNFKLSFISIMAVM 202
Db 380 VAMVVGFAETVVELLKEHSI-LMDEINDIRIIGAITVVI-----LLGISVAGM 428

203 SFTVATIGIGLAIATVAGK--VG-----KTSMTGTAIVGVDTAAOKIW 244
429 EWEAKQIVLLVILLALIGFVIGTPIPLESKPKGPFQYKSEINENFGDPFREBEFF 488
245 RSFOAVGDIAPAVYATVLIETDTRSSPAENKAMKASLVGSTTFFPI- LCG 299
489 SVF-----AIFPAAAGIAGANISGDLADPOSAIPKGTLLAIIITLVVYGIASVGS 542
300 CIGVAAFGN-----NAPG-----DPLT-----DEGFPEPMLIDFAN----- 331
543 CVVADATGNNVDTIVTELNCTSAACKLNDFSSCSSPCSTGLMNNFQVMSWISGFTPL 602
332 -----ACIAVHLIGAYOVF-----AQPFPQVEKKCNKNYPDNK----- 365
603 ISAGIFSATLSSALASIVSAPKIFQALCKDNTYPAFQMPAKGYGKNNPELRGYILFLIA 662
366 ---PITSEYSNNVPELKENISLPRLV-----WRTAVVITTVVAMIPPEF 408
663 LGFTILAEVNIADPIINSEFLASALINFSVFHASLAKSPGRRPG-----FKTY 711
409 NALIGLIGA-----ASFWP--ITVFPVEMHIAOTRIKKYSARW-IALKTMCYVCL 456
712 NMWISLIGALICCIWMEVIMWMAALITVYVILGLYIYV-YKPPVNWGSSFOALTY--- 767
457 IVSLLAAAGSIAGLISSVKTYKP 479
768 -LNALQHSIRLSGVEDHVKQFRP 789

RESULT 14
US-09-895-913A-268
; Sequence 268, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1el Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 268
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-268

Query Match 4.0%; Score 101.5; DB 9; Length 551;
Best Local Similarity 18.9%; Pred. No. 0.12; Indels 175; Gaps 23;
Matches 98; Conservative 75; Mismatches 170; Indels 175; Gaps 23;
20 YTVADPTKNNVDEDEKRGKTWTATASNIITAVIGSVLSLMAIAQI-GWIACTISILLI 78
5 HQIDPPLGNI-----MLSA-----LVALLPILFLSLMFWFKLGYYA----- 42
79 FSPITVFTSTMLAD-CYRAPDPVYTKKNTYMDVWVSYSYGGKRVQLCGVAYQGNLIGV-T 136
43 -AFLSVALSAVIAVLVTKMPVSWGS-----SFL-----YGLFYGLMP 79
137 VGYYITIASISLVAAKSNCPHDKGHTADCTISNYPMAVFGIIOVLQIENFHKLSFLS 196
80 IAWIIIAIFLYKLSV-----SGY-----FEILKSQVOSITLDRHILVIL 120
197 IMAAVMSFTVATIGIGLAIATVAGKVG-----KTSMTGTAIVGVDTAA 240

121 IGFCGSLBEGALIGFGSPRIATTAAILVGLSPLVSAGLCJLANTAPAFAGVGLPISA- 179
241 OKIWRSPQAVGDIAPAVYATVLIETDTRSSPAENKAMKASLVGSTTFFPIILGCC 300
180 -----MASAVGPALII-----SMTGKILFFVSLVPEFYFLM 214
301 IGVAFGNNAAGDPLTFCGFPEPMLID-----FANACIAVHLIGAYOVAQF--I 349
215 DGFQIKETPPAVFIAPFSAQOPLSSNYVLGEBELPGIISALVSLVATALLKFMQPKAI 274
350 PQVPEKKCNKNYPDNKFTSEYSVNNPFLGKNISLPRLWRTAVVITTVVAMIPPEF 409
275 FRSDGKAAS-----FTKSNH-----HICKIYVAMPFVILVILVIMIOPEFK 317
410 AIL-----GLIGASFW-----PL-----TVFPVEMHIAOTRIKKYSARWIA 447
318 ALFERDGLIATSNPFEFFNNISNHIKSPFPVEANQSVSPFVYF-----KELL 365
448 LKTM---CYVCLIVSLLAAAGSIAGLISV-KTYKPR 481
366 INTVGTISIFLAALVSMVLKRVSDALSVFGETLKEMR 403

RESULT 15
US-09-817-514A-2
; Sequence 2, Application US/09817514A
; Patent No. US20020078478A1
; GENERAL INFORMATION:
; APPLICANT: French-Constant, Richard
; APPLICANT: Bowen, David
; APPLICANT: Rocheleau, Thomas
; APPLICANT: Meterfield, Nicholas
; TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS
; FILE REFERENCE: 61645
; CURRENT APPLICATION NUMBER: US/09/817,514A
; CURRENT FILING DATE: 2000-03-26
; PRIOR APPLICATION NUMBER: US 60/191806
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 2516
; TYPE: PRT
; ORGANISM: Photorhabdus luminescens
US-09-817-514A-2

Query Match 3.8%; Score 96; DB 10; Length 2516;
Best Local Similarity 25.1%; Pred. No. 3; Indels 50; Gaps 10;
Matches 50; Conservative 32; Mismatches 67; Indels 50; Gaps 10;
238 TAAQKIMRSFOAVGDIAPAVYATV--LIETD---TLRSSPAENKAMKASLVGSTT 291
183 TKVMEIMSTFPPSGATPYHDAYENVREVIQLDPGLGQLNAPALAGLMHQASLIGIAS 242
292 ---TFEYILGCGCYAFAFGNNAAGDPLTDCGFPEPMLIDFANACIAVHLIGAYOVAQF 348
243 ISPELFINILTEI-----TEGNAEELYKKNFNIPE-----ASLAMEYLLKRYNLSDEE 292
349 IFQVPEKKCN--RNYPDNKF-----ITSEYSVNV-----PFLGKF 381
293 LSQFQKASNFGQOQESYNNOLITPVNSSDGTIVKYRITREYTNAYOMDELFPFGE- 351
382 NISL---FRLVWRTAVYVI 397
352 NYRDLYKFKPNYASYSI 370

Search completed: November 22, 2002, 13:32:10
Job time : 15 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2002, 13:26:12 ; Search time 14 Seconds

(without alignments)
1436.859 Million cell updates/sec

Title: US-09-854-562-2

Perfect score: 2519
Sequence: 1 MSAFTEGHNHSTASGDAY.....SIAGLISSVXTYKPRTHME 485

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	208	8.3	713	1 YNKL1_YEAST	P50944 saccharomyc
2	207.5	8.2	470	1 MTR_NEUCR	P38680 neurospora
3	189.5	7.5	692	1 YK06_YEAST	P26062 saccharomyc
4	167.5	6.6	602	1 YXJ1_YEAST	P47082 saccharomyc
5	165.5	6.6	466	1 UN47_YEAST	P34579 caenorhabdi
6	145	5.8	460	1 YMO2_CAEEL	P34479 caenorhabdi
7	142	5.6	656	1 YAN9_SCHPO	Q10074 schizosacch
8	131	5.2	490	1 YI18_YEAST	P40501 saccharomyc
9	128.5	5.1	565	1 YD08_MYCPV	P35472 mycoplasma
10	125.5	5.0	480	1 YEH4_YEAST	P39981 saccharomyc
11	121.5	4.8	639	1 FU11_YEAST	P38196 saccharomyc
12	121	4.8	591	1 YF7A_SCHPO	Q30818 schizosacch
13	120	4.8	440	1 YJ99_MYCTU	P39589 mycobacteri
14	118.5	4.7	516	1 YWCA_BACSV	P39589 bacillus su
15	115.5	4.6	615	1 YSPK_CAEEL	Q19425 caenorhabdi
16	114.5	4.5	422	1 EXUT_BACSV	O34456 bacillus su
17	114	4.5	1205	1 NKCI_MOUSE	P55012 mus musculu
18	108	4.3	1212	1 NKCI_HUMAN	P55011 homo sapien
19	107	4.2	1023	1 TSCC_DSEAM	P55011 pseudopleur
20	106.5	4.2	490	1 YW0B_BACSV	P49475 bacillus su
21	105.5	4.2	388	1 HCAT_HAENI	P44623 haemophilus
22	105.5	4.2	1095	1 NK02_MOUSE	P55014 mus musculu
23	105	4.2	481	1 Y389_CLOPE	P10818 clostridium
24	105	4.2	657	1 NUOL_RICCN	Q92997 rickettsia
25	104.5	4.1	438	1 PBUX_BACSV	P42086 bacillus su
26	104	4.1	763	1 RGT2_YEAST	Q12086 bacillus su
27	102.5	4.1	430	1 YBAT_ECOLI	P77400 escherichia
28	102.5	4.1	1095	1 NK02_RAT	P55012 ratulus novy
29	102	4.0	705	1 NU5C_HORVU	Q34440 rardus vul
30	101.5	4.0	448	1 GNTF_BACLI	P46833 bacillus li
31	101.5	4.0	580	1 ISPS_SCHPO	P40901 schizosacch
32	101.5	4.0	633	1 FUR4_YEAST	P03316 saccharomyc
33	101	4.0	465	1 ALST_BACSV	Q45066 bacillus su

34	100.5	4.0	488	1 SECY_HALVO	Q977v3 halobacteri
35	100.5	4.0	702	1 NU5C_SORBI	Q33066 sorghum bic
36	100.5	4.0	738	1 NU5C_MAIZE	P46620 zea mays (m
37	100.5	4.0	1191	1 NK01_SQUAC	P55013 squallus aca
38	99.5	3.9	407	1 RFC_SALTY	P26479 salmonella
39	99	3.9	608	1 DSD8_RALSO	Q8xv41 ralteonila s
40	99	3.9	744	1 NU5C_GERJA	P51100 gerbera jam
41	99	3.9	809	1 CICH_TORCA	P35522 torpeda cal
42	98.5	3.9	426	1 VGR0_CHLPN	Q927m4 chlamydia p
43	98.5	3.9	1581	1 VGR0_BEV	P23052 berne virus
44	98	3.9	440	1 YHFE_ECOLI	P37643 escherichia
45	98	3.9	608	1 DIP5_YEAST	P53388 saccharomyc

ALIGNMENTS

RESULT 1
YNKL1_YEAST
ID YNKL1_YEAST STANDARD; PRT; 713 AA.
AC P50944;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 80.0 kDa protein in POL1-RAS2 intergenic region.
GN YN101W OR N2185.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / FY1679;
RX MEDLINE=9626765; PubMed=8701612;
RA Saliz U.E., Bultrago M.J., Soler A., del Rey F., Revuelta J.L.;
RT "The sequence of a 21.3 kb DNA fragment from the left arm of yeast
RT chromosome XIV reveals LEU4, MET4, POL1, RAS2, and six new open
RT reading frames.";
RL Yeast 12:403-409 (1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO YEAST YKL146W AND S. POMBE SPAC31.09C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; Z50161; CA90525.1; -;
CC EMBL; Z71377; CA95977.1; -;
CC SGD; S0005045; YN101W.
CC InterPro; IPR002422; AA/re1_primease2.
CC Pfam; PF01490; Aa trans; 1.
CC Hypothetical protein: Transmembrane.
FT TRANSMEM 243 263 POTENTIAL.
FT TRANSMEM 302 322 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 374 394 POTENTIAL.
FT TRANSMEM 411 431 POTENTIAL.
FT TRANSMEM 439 459 POTENTIAL.
FT TRANSMEM 484 504 POTENTIAL.
FT TRANSMEM 516 536 POTENTIAL.
FT TRANSMEM 562 582 POTENTIAL.
FT TRANSMEM 622 642 POTENTIAL.
FT TRANSMEM 649 669 POTENTIAL.
FT TRANSMEM 693 713 POTENTIAL.
SQ SEQUENCE 713 AA; 80025 MW; EAB458A2FEB660FB CRC64;
Query Match 8.3%; Score 208; DB 1; Length 713;
Best Local Similarity 21.4%; Pred. No. 1,8e-07;
Matches 116; Conservative 197; Indels 162; Gaps 22;

```
CC      5 NTEGHNSTAESGDAYTVSDPTKNVDSD----- 32
CC      : : : : :
CC      207 NDSASSDFTSHESDSINQSSPSSNQDIDKVPFLTRNFLEFLYVGHGFASESFEODDFIPDS 266
CC      : : : : :
CC      33 -----GREKPT-----GTLWTASAHII--TAVIGSGVLSLAWIAQLG 68
CC      : : : : :
CC      267 SNMIRGEDSALLSPDKMVKLPKAKGTSTKKVFLILLKSPFTGTGVLFLPNAFHNGG 326
CC      : : : : :
CC      69 WIAGTSILLIFSTYFTSTMLADCVRAPDPVTGKKNYTYMDVRSYLGGRKVQLCGVAQ 128
CC      : : : : :
CC      327 LFFSVSMALPFGIYSW-----CY-----YILVQAKSSCGVSS 359
CC      : : : : :
CC      129 YGN-----LIGTVGYCTITASISLVANGKS-----NCFHDKGHTADCTIS 168
CC      : : : : :
CC      360 FEDIGLKLGPWMKILLFLSVITQVSGFAGYMIPTAKNLQAFLDNVFH-----VGVL 412
CC      : : : : :
CC      169 NPYMAVE-GIIQVILSQIPNFHKLSPFLSMAAVMSFTYATIGIGLAIAITVAGGKVGKTS 227
CC      : : : : :
CC      413 PLSYLVWFQTIIFPLSPFIRNISKLSLPSLLANFF-----IMAGLVIVLIPTAKRLPFD 466
CC      : : : : :
CC      228 MTGT-AYGVQDVTAQKTRWRFQVGDIAFAYAVATVLEIQTDLTLESSPAENKMKRASLV 286
CC      : : : : :
CC      467 LMGTPAMGVVGLNADRWTLF--IGTAIFAFEGIGLIIPVQDSMRN---PEKFLVLVALV 521
CC      : : : : :
CC      287 GVSTTTFFYLCCGIGYAAGFNAPGDFLTDFGFFPEPFLIDPANACIAVHLIGAY---- 342
CC      : : : : :
CC      522 -ILTATILFISIALGLVAYGSNV-----QTVILLNLPQSNFVNLIQLFYSLA 569
CC      : : : : :
CC      343 -----QVFAQPIQFQVEK-----KCNRYPNKFTTSEYVNVPLGKFNISLPLVWR 391
CC      : : : : :
CC      570 IMLSTPQLQF--PAIKTIENKFPKFTKIVRHDDUTTRVELR-PNSKLM---WKIKWL 623
CC      : : : : :
CC      392 TAYV--VITTVVAMIFPF-----FNAILGLIGAASFVPLTVPPVEMHIAQTKIKKYSARW 445
CC      : : : : :
CC      624 KNFIRSIIVIVWSIAYFGSDNLDKDFSVIGSLACIPLVVIYPSMLHRLGRNSLPETKGEF 683
CC      : : : : :
CC      446 IALKTMCVCLLI 547
CC      : : : : :
CC      684 WRFKPLDITILI 695
CC      : : : : :

RESULT 2
MTR_NEUCR
ID MTR_NEUCR STANDARD; PRT; 470 AA.
AC P3680;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE N amino acid transport system protein (Methyltrophophan resistance
DE protein).
GN MTR.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariatales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oak Ridge;
RX MEDLINE=95095055; PubMed=8001794;
RA Dillon D.; Stadler D.;
RT "Spontaneous mutation at the mtr locus in neurospora: the molecular
RT spectrum in wild-type and a mutator strain."
RL Genetics 138:161-74 (1994).
RN [2]
RP SEQUENCE OF 210-470 FROM N.A.
RX MEDLINE=92146948; PubMed=1838345;
RA Koo K.; Stuart W.D.;
RT "Sequence and structure of mtr, an amino acid transport gene of
RT Neurospora crassa."
RL Genome 34:644-651 (1991).
CC -!- FUNCTION: REQUIRED FOR THE TRANSPORT OF NEUTRAL ALIPHATIC AND
CC AROMATIC AMINO ACIDS VIA THE N SYSTEM.
```

```
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; L34605; AAA33600.1; -.
CC      EMBL; S81767; AAB21410.1; -.
CC      PIR; A54551; A54551.
CC      InterPro: IPR002422; AA/re_l_primease2.
CC      Pfam; PF01490; Aa_trans; 1.
CC      Transport; Amino-acid transport; Transmembrane.
CC      DOMAIN 1 56 EXTRACELLULAR (POTENTIAL).
CC      TRANSMEM 57 77 POTENTIAL.
CC      TRANSMEM 78 98 POTENTIAL.
CC      DOMAIN 99 131 EXTRACELLULAR (POTENTIAL).
CC      TRANSMEM 132 152 POTENTIAL.
CC      DOMAIN 153 168 CYTOPLASMIC (POTENTIAL).
CC      TRANSMEM 169 189 POTENTIAL.
CC      TRANSMEM 191 211 POTENTIAL.
CC      DOMAIN 212 236 CYTOPLASMIC (POTENTIAL).
CC      TRANSMEM 237 257 POTENTIAL.
CC      DOMAIN 258 275 EXTRACELLULAR (POTENTIAL).
CC      TRANSMEM 276 296 POTENTIAL.
CC      DOMAIN 297 316 CYTOPLASMIC (POTENTIAL).
CC      TRANSMEM 317 337 POTENTIAL.
CC      DOMAIN 338 357 EXTRACELLULAR (POTENTIAL).
CC      TRANSMEM 358 378 POTENTIAL.
CC      DOMAIN 379 386 CYTOPLASMIC (POTENTIAL).
CC      TRANSMEM 387 407 POTENTIAL.
CC      DOMAIN 408 427 EXTRACELLULAR (POTENTIAL).
CC      TRANSMEM 428 448 POTENTIAL.
CC      DOMAIN 449 470 CYTOPLASMIC (POTENTIAL).
CC      SEQUENCE 470 AA; 51162 MW; E8132D1A62373300 CRC64;

Query Match 8.2%; Score 207.5; DB 1; Length 470;
Best Local Similarity 22.3%; Pred. No. 1.2e-07;
Matches 110; Conservative 75; Mismatches 192; Indels 117; Gaps 24;

QY 31 EDGREK--RTGTWLTASAHIIITAVIGSGVLSLAWIAQLGAGTSILLIFSPFTYFTST 88
DB : : : : :
QY 89 MLADCYRAPDPVTGKKNYTYMDVRSYLGGRKVQLCGVAQYGNLIGVTGVTITASISLV 148
DB : : : : :
QY 103 VIGQTKLKHPEIA-----HYADVGRVMFGRMGVEIISPMFVLQILFIVGSHVLTGTTMMWG 157
QY 149 AVGKSNCFHDKGHTADCTTISNYPYMAVFGIIQVI-----LSQIPNFHK-----LSFLSIM 198
DB : : : : :
QY 158 TI-----TONGN-GTCSL-----VGIYSAILFLFLAIPPSFAEVALIGIDFVSIC 203
QY 199 AAVMSFTYATIGIGLAIAITVAGGKVGKTSMTGTAVGV-----DVTAAQKTRWRFQAV 250
DB : : : : :
QY 204 AAIL-----ITMIATG--IRSSHQEGGLAAVPMSCWPKEDLSLAE---GFIIV 246
QY 251 GDIAFAYAYATVLIETIQTDLRSPSPAEKAKRASLVGVSTTFFYILCGGIGYAAGFNNA 310
DB : : : : :
QY 247 SNIVFAYSFAMCQSPFMDHPTSDYKKSIVALGLIEI-----FIYVTGGVYAFVGPV 302
QY 311 PGDFLTD-----FGFEFFWLIDPANACIAVHLIGAYOVFAQPIFOFVEKKCNRY 361
DB : : : : :
QY 303 QSPALLSAGPLLAKVAFGIAPL---VIFSGSINTVVVSRYL-----ERIW 346
QY 362 PDNKFITSEYSNVNP-----FLGKFNISLFRVWRTAYVITTVVAMIFPFNAILGLI- 415
DB : : : : :
QY 347 PNN---VIRY-VNTPAGVMVNLG-PDFGITLIW-----VIAEAIFFSLLAICS 392
QY 416 -----GAASFVPLTVPPVEMHIAQTKIKKYSARWIAKTMCV-----CLIVSLAAAGS 466
DB : : : : :
```

Db 393 ALPISGSEFFYPALMYKTRNDKASQKKYFLD--ALNMLCFVIGMIGITGYAAID 450
 QY 467 IAGLISSVKTYYKPF 480
 Db 451 IMDRYDHGKVSKEP 464

RESULT 3
 YK06 YEAST
 ID YK06 YEAST STANDARD; PRT; 692 AA.
 AC P36052;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hypothetical 75.5 kDa protein in SDH1-CIM5/YTAJ intergenic region.
 GN YK146W OR YK1600.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=94378720; PubMed=8091859;
 RA Vandenbol M., Bolle P.-A., Dion C., Portetelle D., Hliger F.;
 RT "DNA sequencing of a 36.2 kb fragment located between the FAS1 and
 RL LAB loci of chromosome XI of Saccharomyces cerevisiae.";
 RN [2]
 RP SEQUENCE OF 447-692 FROM N.A.
 RA Rad M.R., Xu G., Kirchbach L., Fritz C., Keuchel H., Hollenberg C.P.;
 RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO YEAST YNL01W AND S.POMER SPAC3H1.09C.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC
 DR EMBL; Z26877; CAAB1508.1; -
 DR EMBL; Z28146; CAAB1988.1; -
 DR PIR; S37976; S37976.
 DR SGD; S0001629; YK146W.
 DR InterPro: IPR002422; AA/re1_primease2.
 DR Pfam: PF01490; Aa.trans; 1.
 DR Hypothetical protein, Transmembrane.
 KM DOKAIN 261 276 POT-GLU.
 FT TRANSSEM 302 322 POTENTIAL.
 FT TRANSSEM 329 349 POTENTIAL.
 FT TRANSSEM 374 394 POTENTIAL.
 FT TRANSSEM 412 432 POTENTIAL.
 FT TRANSSEM 443 463 POTENTIAL.
 FT TRANSSEM 483 503 POTENTIAL.
 FT TRANSSEM 519 539 POTENTIAL.
 FT TRANSSEM 561 581 POTENTIAL.
 FT TRANSSEM 607 627 POTENTIAL.
 FT TRANSSEM 630 650 POTENTIAL.
 FT TRANSSEM 665 685 POTENTIAL.
 FT TRANSSEM 692 AA; 75459 MW; 379054D69094A0FS CRC64;
 SQ SEQUENCE

Query March 7.5%; Score 189.5; DB 1; Length 692;
 Best Local Similarity 23.1%; Pred. No. 3.5e-06;
 Matches 106; Conservative 73; Mismatches 183; Indels 97; Gaps 18;

QY 8 GHNHSTASGDAYTVSDPTXNV-----DEGKERKRTGWTLTASAHITTAIVIGSGV 57
 Db 256 GEDLSSEEESEEESEEESEEEALTESTVLVSREHGRHPHKSTVAVALLKSFVGTGV 315
 QY 58 LSLAWALNQDLINGTSLILIFSPITY--FTSTMADCYRAPDPVTKGNKNTYVMVWVASY 115

Db 316 LFLPKAFNHNGWGFSAALICLLSCALISYGGFVSLI-----TTKDKV--- 355
 QY 116 LGRKVLQCGVAOYGN-----LIGVYVGTITNLSLVAVGKS-----NC 155
 Db 356 -----GVDDYGDGMGRILVGPKNKFAISSIALSQIGPSAAVYFTATNLQVESENF 406
 QY 156 FHDKHTADCTISNPFYMAVFGIIQVILSQIPNFHKLSPFLSMAAVSFYATIGLAI 215
 Db 407 FHLK--PGSISLTYFAQV--LIFVPLSTRNIAKLSGTALIDL-----ILGLVY 456
 QY 216 ATTAGKVKGTSMGTGAVGVDTAAKTIWRSFOAVGDIAPVATVLEIQDTLRSSPA 275
 Db 457 VYVY--SIYVIANGVASDTMTMFNAMSLE--IGTAIFTEEGIGLIPQESMK----- 508
 QY 276 ENKAKRASLVGV-STTTFEYILCGICGYAAGFNNAPOGFLTDEGFEPFELIDFANACI 334
 Db 509 -HPGHFPPSLSAMCVIANVIFISGLLCCAAAGSVKVVVLNPPDQTSYLVTVQLLYAL 567
 QY 335 AVHLIGAVQVPAQPIQFQPEKKNRYPDNKRTSEYVNVVPIGKRNISLRLVMTAY 394
 Db 568 AILSTPLQLE--PAIRLE--NMTFPSN--ASGVNPKVKWLKNV-----FRCAI 612
 QY 395 VVITTVVAMI--PEFPNALIGLGAASFVFLTVYFPEVEMH 432
 Db 613 VVLSTIAWGANDDKRFVLSVGSFACIPLIYIYRPLH 651

RESULT 4
 YKX1 YEAST
 ID YKX1 YEAST STANDARD; PRT; 602 AA.
 AC P47082;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Hypothetical 65.3 kDa protein in FRE3-MPP10 intergenic region.
 GN YK001W OR J1409 OR YOR83.4.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / FY1679;
 RA de Haan M., Smits P.H.M., Grivell L.A.;
 RL Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO C. ELBEGANS UNC-47.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC
 DR EMBL; X87611; CAAB0922.1; -
 DR EMBL; Z49501; CAAB9523.1; -
 DR SGD; S0003761; YOR001W.
 DR InterPro: IPR002422; AA/re1_primease2.
 DR Pfam: PF01490; Aa.trans; 1.
 DR Hypothetical protein, Transmembrane.
 KM TRANSSEM 220 240 POTENTIAL.
 FT TRANSSEM 241 261 POTENTIAL.
 FT TRANSSEM 287 307 POTENTIAL.
 FT TRANSSEM 322 342 POTENTIAL.
 FT TRANSSEM 345 365 POTENTIAL.
 FT TRANSSEM 390 410 POTENTIAL.
 FT TRANSSEM 430 450 POTENTIAL.
 FT TRANSSEM 467 487 POTENTIAL.
 FT TRANSSEM 518 538 POTENTIAL.
 FT TRANSSEM 544 564 POTENTIAL.

```
FT TRANSMEM 581 601 POTENTIAL.
SQ SEQUENCE 602 AA; 65345 MW; C2F321E76C21C2B4 CRC64;

Query Match 6.6%; Score 167.5; DB 1; Length 602;
Best Local Similarity 19.6%; Pred. No. 0.00012;
Matches 99; Conservative 92; Mismatches 208; Indels 105; Gaps 19;

QY 17 GDYVTSPTKNVDE-----DGRKRTGTWLTASA----- 46
DB 156 GDSGILRPTASLTSMSSGGAGRRFTNNDMSIVVRVGVGDKVVVTLAQSTAPQTF 215
QY 47 HIITAVIGSVLSLAWIAQOLGIAQIAGTSTIILIPSFYITSTMLADCYRADPVPVTKRNY 106
DB 216 NSINVIGIGLALPLGLKAGVIGITLMAIFALATFCTALLSRCLDT-DPTL----I 270
QY 107 TYMDVVRVYLGGRKVLQCGVAQVGNLIGVTVGTTITASISLVAVGSKNCFHDKGHTADCT 166
DB 271 SYADLGAAFGTKRALISALFTDLGLSGVSLVILFGDSLNAL----- 314
QY 167 ISNVPYMAV--FGIIQVILSQINPHKLSPLS--INAAVMSPTYATIGIGLAIATVAGGK 222
DB 315 ---FPQYSTTFPKIVSFPIVTPVPLSVLSNISLGLS-----TTGTVLVI CCG--- 364
QY 223 VOKTSMGTAVGVDTAAQKIW-----RSFQAVGDIAFAYAYATVLIIBIQDILRSSPAEN 277
DB 365 LYKSSSPGSLNPMETS---MWPIDLKHLCISIGLSACMGHGVFPNKLKTDMRHPDKPK 421
QY 278 KAMKR---ASLVGVSTTTPFYILCGGICGYAAFGNNAPGDFLTDFGFE--PFWLIDFAN 331
DB 422 DCLKTYKITSTVDIGTAV-----IGFLMEGNLVKDEITKNVLLTEGYPKFVYGLIS 473
QY 332 ACIAVHLIGAVQVFAQIPQFVEKKCNRNYPDNKFTSYSVNVPFLGKFNISLFLVWR 391
DB 474 ALMTIPIAKTPLNAREPIVSLVDVLMNVQHID-----EASAIKRAAKGLQVFNRFI 527
QY 392 TAYVVITTVVAMIFPFFNAILGLIGASFWPLTVYFP--VENHIAQTIKIKYSARWIALK 449
DB 528 N---VVFVLIAINPEFDKIIAFLGACLCFTICLIILPCWFVYLRCLKTKTIKWE-----R 578
QY 450 TWCYVCLIVS--LLAAGSIAGLIS 472
DB 579 VACHVTICISVVLSTGLVGAAIIS 602

RESULT 5
UN47 CAEL
ID UN47 CAEL STANDARD; PRT; 486 AA.
AC P34579; O17475;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Uncoordinated protein 47 (Protein unc-47).
GN UNC-47 OR T20G5.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98007977; PubMed=9349821;
RA McIntire S.L., Reimer R.J., Schuske K., Edwards R.H., Jorgensen E.M.;
RT "Identification and characterization of the vesicular GABA
RT transporter";
RL Nature 389:870-876 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Berks M., Smith A.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
```

```
CC -- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -- SIMILARITY: TO YEAST YJR001W.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; AF031935; AAB87066.1; --
CC EMBL; Z30423; CAA83006.2; --
CC PIR; S42372; S42372.
CC WormPep; T20G5.6; CE25119.
CC InterPro; IPR002422; AA_rel_primease2.
CC Pfam; PF01490; Aa_trans; 1.
CC Transmembrane.
CC FT TRANSMEM 94 114 POTENTIAL.
CC FT TRANSMEM 120 140 POTENTIAL.
CC FT TRANSMEM 168 188 POTENTIAL.
CC FT TRANSMEM 204 224 POTENTIAL.
CC FT TRANSMEM 229 249 POTENTIAL.
CC FT TRANSMEM 264 284 POTENTIAL.
CC FT TRANSMEM 306 326 POTENTIAL.
CC FT TRANSMEM 342 362 POTENTIAL.
CC FT TRANSMEM 399 419 POTENTIAL.
CC FT TRANSMEM 422 442 POTENTIAL.
CC FT TRANSMEM 458 478 POTENTIAL.
CC SQ SEQUENCE 486 AA; 54439 MW; 6435C3B338552925 CRC64;

Query Match 6.6%; Score 165.5; DB 1; Length 486;
Best Local Similarity 19.4%; Pred. No. 0.00012;
Matches 97; Conservative 91; Mismatches 183; Indels 129; Gaps 22;

QY 13 TAESG-----DAYTVSDPTK---NVDEDRKRTG-----TWLTASAHIT 50
DB 41 TGESGPPPHDRLEPIQESVVSQPKDDINKQEEAKDDGGEASEPISALQANVTN 100
QY 51 AVTSGVLSLAWIAQOLGIAQIAGTSTIILIPSFYITSTMLADCYRADPVPVTKRNYTVM 109
DB 101 AIOGMFIVGLPIAVKVGWWS--IGAMVGVAIVCVYVWTVGLLIECLYE----NGVKRKTYR 155
QY 110 DVRSYLGGRKVLQCGVAQVGNLIGVTVGYTITASISLVAVGSKNCFH--DKGHTADCTI 167
DB 156 EIADFYKPGFGKWVL-AAQLTELLSTCIYLVLAADLL-----QSCFSPVDKAGWMTS 209
QY 168 SNPYMAVFCIIQVILSQINPHKLSPLS-----MAVMSFTYATIGIGLAIATVA 219
DB 210 ASLLTCSFLDDLQIV-SRLSFFNAISHLIVNLIMVLYCLSFVSQMSFTTTFSLNINTLP 268
QY 220 GKGVGKTSMTGTAVGVDTAAQKIWRSFQAVGDIAFAYAYATVLIIBIQDILRSSPAENKA 279
DB 269 -----TIGVVVFGYTSHIFLPNLEGNMK-NPAQFNV 299
QY 280 MKRASLVGVSTTTFVILCGCIGYAAFG-----NNAPGDFLTDFGPEPFWLIDFAN 331
DB 300 MLKWSHTAAAV---FKVFMGLGFLTFGELTQEISNSLEN-----QSFKIL--VN 345
QY 332 ACTAVHLIGAVQVFAQIPQFVEKKCNRNYPDNKFTSYSVNVPFLGKFNISLFLVWR 391
DB 346 LILVVKALLSYPLPFAVAVQLKNNLFLGYPOTPP--TSCYSPD-KSLRENAVTL-----R 398
QY 392 TAYVVITTVVAMIFPFFNAILGLIGASFWPLTVYFPVEMHIAQTKIKKYSARWIALKTM 451
DB 399 IILVLFRLFVALSVPYLVLMGLVGNITGTMLSFITWLPALFHL-----YIKEKTL 447
QY 452 -----CYVCL 456
DB 448 NNPEKRFQGIIMGCVCI 467

RESULT 6
```


YMW2_CABEL STANDARD; PRT: 460 AA.

AC P34479; O9UJEO.

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein F59B2.2 in chromosome III.

GN F59B2.2.

OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

OC Rhabditidae; Pelodicerinae; *Caenorhabditis*.

OX NCBI_TaxID=6239;

RA SEQUENCE FROM N.A.

RP STRAIN=Bristol N2;

RC MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A., Fulton L., Gardner A., Green P., Hawkes T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Latster N., Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R., Sims M., Smaldon N., Smith A., Smith M., Sonnenhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sprat J., Wodham P.;

RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*."

RT Nature 368:32-38(1994).

RL [2]

RN REVISIONS.

RA Durbin R.;

CC Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ch/announce/> or send an email to license@ebi-sb.ch).

CC -----

CC EMBL; Z11505; CAA77582.1; -

DR PIR; S31123; S31123.

DR WormPep; F59B2.2; CE20899.

DR InterPro; IPR002422; AA/rel_primease2.

DR Pfam; PF01490; AaTrans; 1.

DR Hypothetical protein; Transmembrane.

KM TRANSMEM 64

FT TRANSMEM 132 152 POTENTIAL.

FT TRANSMEM 172 192 POTENTIAL.

FT TRANSMEM 194 214 POTENTIAL.

FT TRANSMEM 236 256 POTENTIAL.

FT TRANSMEM 270 290 POTENTIAL.

FT TRANSMEM 316 336 POTENTIAL.

FT TRANSMEM 362 382 POTENTIAL.

FT TRANSMEM 383 403 POTENTIAL.

FT TRANSMEM 426 446 POTENTIAL.

SO SEQUENCE 460 AA; 51338 MW; 0D7809A9A4E15BA CRC64;

Query Match 5.8%; Score 145; DB 1; Length 460;

Best Local Similarity 20.3%; Pred. No. 0.0031;

Matches 109; Conservative 79; Mismatches 186; Indels 164; Gaps 28;

OY 7 EGHNSTAESDAYTSDPTKYNDEGKERKTGWTLPASAHIT--AVTSGVSL--A 61

DB 5 EGRERAVTEDEASNDMDGRALVQPPA---RSGDVITTPRAVLTUSKSFNAGCSLEPYA 61

OY 62 WAIADGIAGTSLILFSPFITVSTMLADCYRAPPDVTGKKNVTYMDVARSYLGGRKV 121

DB 62 WTLGGL-WVS-----FVMSFV----- 76

OY 122 QLGAGVAGNLTIGVTVGTITASISLVAVGKSNCFHDKHTAD--CTISNVPW-----A 174

DB 77 -TAGLNMYGNHILVRSASQHLAKKSDRSAL-----DYGHAKKVCYSDIRFLRNNSKA 128

OY 175 VFGIIQVILISQIPNFHKLSPFLSIIAAVMSFYATIGIS-----LAIVATGAKV 223

DB 129 VMVEFVNTIL-----FYLGWCVAIIIFISDNLVNL-VGDHIGSTRHOQMILMAVSILFI 183

OY 224 GKISM-----TGTNGVDTAAQ-KIMRSQA-----VGD1 253

DB 184 LUTNMFTEKATVSEFFLVSSVFVIGAAVIMQYTOQNQMDKLPALNFTGTTMIGMS 243

OY 254 AFAYAVATVLEIETDTRSSPAENKRAKSLVSGT-----TFPYILCCIGYAAFGNN 309

DB 244 MYAREGQTMILPIENLT-----DNPAAFLAPGVSTMTMICTFMALGFFGYTGFDS 298

OY 310 ARGDFLDF---GFEPF-----FWLIDFNAACIAYVLIQAYQVAFQPFQPEKKCNVNP 362

DB 299 IAPITTVNVEKGLSTVNFVLMLOSLGNSIAMTVV--IDMF-----FNGRRKFGARFP 352

OY 363 DNKEITSESVNVP-FLGKFNISLPELVWRTAVVITTVVAMIFPFENAILGLIGAASF 421

DB 353 -----NPKMLSDKG--FRVFW-----VLVTYLMVLIPLKLEIMIPLVGVSQA 394

OY 422 PLTYRFP--VEHMIQTKIK--KYCARPIAKMVCYCLVSLAAGSIAGLISVS 474

DB 395 LCAIIFPFPFEMITFTWMDKGLTYROR---WTKIFNLVMAIGVFAIIAGVYTN 448

RESULT 7

YAN9_SCHPO STANDARD; PRT: 656 AA.

ID 010074;

AC 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein C3H1.09c in chromosome I.

GN SPAC3H1.09C.

OS Schizosaccharomyces pombe (Pisaion Yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RA SEQUENCE FROM N.A.

RP STRAIN=972;

RC MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Beat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins S., Connor R., Cronin A., Davis P., Feltwell J., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Sissons M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Voicakeert G., Aert R., Robben J., Grymopritz B., Weltjens I., Vanstreels B., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Medler H., Mambutt R., Purnelle B., Goffeau A., Cadieu E., Deyano S., Gloux S., Lelaune V., Mortier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J., RA Shpakovski G.V., Uesary D., Barrell B.G., Nurse P.;

"The genome sequence of *Schizosaccharomyces pombe*."

```

RL Nature 415:871-880(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO YEAST YKL146W AND YNL101W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z68144; CA92262.1; -.
DR InterPro: IPR002422; AA/rel_pmasee2.
DR Pfam: PF01490; Aa_trans; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 280 300 POTENTIAL.
FT TRANSMEM 307 327 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 389 409 POTENTIAL.
FT TRANSMEM 419 439 POTENTIAL.
FT TRANSMEM 457 477 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
FT TRANSMEM 537 557 POTENTIAL.
FT TRANSMEM 578 598 POTENTIAL.
FT TRANSMEM 601 621 POTENTIAL.
FT TRANSMEM 636 656 POTENTIAL.
SQ SEQUENCE 656 AA; 73060 MW; 7FD1E8F649ED83D2 CRC64;

Query Match 5.6%; Score 142; DB 1; Length 656;
Best Local Similarity 21.1%; Pred. No. 0.0074;
Matches 97; Conservative 71; Mismatches 155; Indels 136; Gaps 22;

QY 48 IITAVIGSVLSLAWAIAQLGWIAGTSILLIFSFITYFTSTMLADCYRAPDVTGKRYNT 107
DB 284 LLKSFVGTGVLFKAPKLGGLVFSSATLLVGLSHICFLLIQ-----T 329

QY 108 YMDVVRVY--LGGKRVQLGVQVAGYGNLGVYGYTTASISLVAVGKSNCF----- 156
DB 330 RMKVPGSGDIG-----GTLGPHMRFAILASIVSQIFSAISVFVASTLQ 377

QY 157 -----HDKGTADCTISNYPYMAVFGIIQ-----VILSQIPNFHKLSPLSMAAV-- 201
DB 378 ACVKVISTHREYH-----LAVETIQFLVFVPLSLVRKISLATALIADVFI 426

QY 202 -MSFTYATIGIGLAIAIVAGKVKTSMTGTAVGVDDVTAQAQIMSFQAGVDIAFAYA 260
DB 427 LGLILYLFWDVITLTKGIADVAMFNKTDPSL-----FIGVAITTEGI 471

QY 261 TVLIEIQDTRSSPAENKAMKASLVGV-STTTFFYLGCIGCYAAGN-----NAP 311
DB 472 CLILPIQEQM---AKPNLPKL-LTGVMAAISLLFISIGLLSYAAGSKVKVTVILNMP 526

QY 312 GDFLTDGFPFPFLIDFANACIAVHLIGAYQVPAQPIQFQVKEKCNRYPNDFITSEY 371
DB 527 ESTPT-----VITQFLYA-TAILSTPLQVF--PAATIEQ-----GIFTR-- 565

QY 372 SVNPFILGKFNISLRLVWRATYV-VITTVAMIPFP-----FNAILGLIGAAFWELTV 425
DB 566 -----GKNR---RKIKRKNVLRVLVILAILISWAGSRDLDFVSMVGSVCCIPLIY 615

QY 426 YFPVEMHIAQTKIKYSARVIALTKCYVCLIVSLIAA 464
DB 616 MYPMLH-----YKACANNWI-LRTLDIFNFTIGAFAMA 648

RESULT 8
YI18 YEAST STANDARD; PRT; 490 AA.
AC P40501;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)

```

```

DE Hypothetical 53.7 kDa protein in SGAL-KTR7 intergenic region.
GN YIL088C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288c / AB972;
RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hanly N., Hornsall T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z46728; CA986706.1; -.
DR SGD: S0001350; YIL088C.
DR InterPro: IPR002422; AA/rel_pmasee2.
DR Pfam: PF01490; Aa_trans; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 7 27 POTENTIAL.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 109 129 POTENTIAL.
FT TRANSMEM 144 164 POTENTIAL.
FT TRANSMEM 191 211 POTENTIAL.
FT TRANSMEM 222 242 POTENTIAL.
FT TRANSMEM 265 285 POTENTIAL.
FT TRANSMEM 398 418 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT TRANSMEM 464 484 POTENTIAL.
SQ SEQUENCE 490 AA; 53736 MW; D3D15CD8253220A1 CRC64;

Query Match 5.2%; Score 131; DB 1; Length 490;
Best Local Similarity 17.1%; Pred. No. 0.033;
Matches 84; Conservative 82; Mismatches 166; Indels 158; Gaps 19;

QY 38 TGTWLTASAHIITAVIGSVLSLAWAIAQLGWIAGTSILLIFSFITYFTSTMLADCYRAP 97
DB 4 TSSALSSTANLVKTIIVGAGTLAIPYSFKSDGLVGVILTLAAVTSGLGLFVLSKSKT- 62

QY 98 DPTVGKEN-----YTYMDVVRVSLGGRKVLQGV-----QYGNLIGVTGVT 140
DB 63 --LIPRNSFFFLCMLTPTTAPIDFLAMIVQCFVGLSVLVLIGDLFFGLFGERNY 120

QY 141 ITASISLVAVGKSNCFHDKGTADCTISNYPYMAVFGIIQVILSQIPNFHKLSPLSMAA 200
DB 121 IIASAVII-----IPLCLVKLDQLKYSSILGLFALAVISILV- 158

QY 201 VMSPTYATIGIGLAIAIVAGKVKTSMTGTAVGVDDVTAQAQIW--RSQAV-----GDI 253
DB 159 ---FSHFVF-----ELGKELT-----NILRNDICWKIHDFFGLLSTFSII 197

QY 254 AFAYATVLIETQDTRSSPAENKAMKASLVGVSTTTFFYLGCIGCYAAGFNAPGD 313
DB 198 IFAPGTSMNLFPMINELKNSMENITFVINNSISLSTALF--LIVGLSGYLTFGNETLGN 255

QY 314 FLTDFGPFEP--FWLIDFANACIAVHLIGAYQVPAQPI----- 349
DB 256 LMLN---YDPSNIWV-ICKFCLGSMILSLPFLPHPLFPLRNVNVIWITTYGANPEE 311

QY 350 -----FQVEKKC---NRNYPDNKFTITSEYSVN 374

```

```

Db 312 DPOVSEYTRASNLRPISTMYEDPADPSDALDNTSYNEQECCLPNQNF-DNDSISSQNNM 370
OY 375 -----VPLFGKPNISLFLVWRTAVYVTT-VVAMIFPFAALLGLIGA 418
Db 371 DERGTMAVADNDENHAFVYKS-----REYWTALLLSMTYALSLVQSFALVLSFVGAT 424
OY 419 SFMPLTVFP 428
Db 425 GSTSISFTLP 434

RESULT 9
YD08_MYCPN STANDARD; PRT; 565 AA.
AC P75472; 008088;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN508 (F10_ORF565).
GN MPN508 OR MP528.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxId=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreid R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97252497; PubMed=9098066;
RA Krause D.C., Proft T., Hedreya C.T., Hilbert H., Plagens H.,
RA Herrmann R.;
RT "Transposon mutagenesis reinforces the correlation between Mycoplasma
RT pneumoniae cytoskeletal protein Hmw2 and cytodifferentiation."
RL J. Bacteriol. 179:2668-2677(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; TO M.PNEUMONIAE MPN095.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; TO M.PNEUMONIAE MPN096.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk.)
CC -----
CC DR EMBL; AEO00052; AAB96176.1; -
CC DR EMBL; U59896; AAB52525.1; -
CC DR InterPro; IPR002293; AA/rel. Primease1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 8 28
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 95 115 POTENTIAL.
FT TRANSMEM 137 157 POTENTIAL.
FT TRANSMEM 167 187 POTENTIAL.
FT TRANSMEM 227 247 POTENTIAL.
FT TRANSMEM 268 288 POTENTIAL.
FT TRANSMEM 314 334 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 424 444 POTENTIAL.
FT TRANSMEM 480 502 POTENTIAL.
FT TRANSMEM 516 536 POTENTIAL.
FT TRANSMEM 565 AA; 62560 MW; ACC745B43D2184E2 CRC64;
SEQUENCE

```

```

Query Match 5.1%; Score 128.5; DB 1; Length 565;
Best Local Similarity 19.5%; Pred. No. 0.057;
Matches 99; Conservative 69; Mismatches 165; Indels 175; Gaps 23;

OY 41 WLDSAHITTAIVAGSGVLSLMAIAQLGWIAGTSTILLIFSTITYFTSTMLADCYRADPV 100
Db 48 WLVASV-----AVIAMAALALIEIASVRNDN-TLSIISVVKFERNRL-----Y 88
OY 101 TGRKNY-TYMDVVSRYIGRKVQLCGVAD-YGNLIGTVG-----YTITASISLVAV 150
Db 89 HGCKNEMTYLPLTFPPPLFVFCISIDQFGRGLGLGTGAHRTVSQDMLMALALI-- 146
OY 151 GKSNCPEHDKHTADCTISNV-----PYNAVFGIIQVILISQIPNFKLSPLSISMAAVSFT 205
Db 147 -----ITTYFLTIPLVAKVGNIONVVASVKFIPVFPDI----- 182
OY 206 YATIGIGLALATAGGV-----GKTSMTGAVGVDT 238
Db 183 -----LGFYVAGNGELKAVKALVOPPOINGATATSPOLVQAGGITRFGTAGNG-- 235
OY 239 AAKIWRSPQAVGDIAFAYAVATVLEIQDTRSSPAENKAKRASLVGSYTFPFIYC 298
Db 236 -----SPFISIAIFPAAVDGFYVTAGLQSEMR-----EPKXTPMALPLGLITITLFPYIL 284
OY 299 G-----CIGYAAFGNNAFGDPLTDFGFPFPLIDFANACIAVHLIGAYQ 343
Db 285 AVALSTINGLPSGMESMG-KLPNKRAGQIV--FG-----VNMMLGIGVLGIIIN 332
OY 344 VPAQPIFQVYK-KCNRRNPDKNETSEYVAVPLGKFNISLFLVWRTAVYVITVVA 402
Db 333 GPALMAAPFVEDLLAQGDLPFMKQVGRILNPKRPVVG-----VYICVLVLTIV 380
OY 403 MIFPFPNAILGLIGAASFPLTYVPPVEMHIAQTIRK-----YS-----ARWMLKT 450
Db 381 QV-----LFTVIGALAYLPTVADY--KNYVTEIDKLNMKMLYSPSDMLATWLTSLPT 431
OY 451 MCYCVLIVSLAAAGSIAGLISSVKTYK 478
Db 432 FAFI-----ACAFGALVVRKTKK 450

RESULT 10
YH4_YEAST STANDARD; PRT; 480 AA.
AC P39981;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 53.3 kDa protein in HXT8-CAN1 intergenic region.
GN YH064C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berrio A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung B., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namach A., Norgren R., Oefner P., Oh C.,
RA Petzel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yellon M., Bortstein D., Davis R.W.;
RT Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RL -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: SOME, TO C.ELEGANS P13H10.3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

```
CC -----
CC EMBL; U18795; AAB65023.1; -.
CC SGD; S0000790; YEL064C.
CC InterPro; IPR002422; AA/rel_pmease2.
CC Hypothetical protein; Transmembrane.
KW TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 95 115 POTENTIAL.
FT TRANSMEM 145 165 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 263 283 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 338 358 POTENTIAL.
FT TRANSMEM 394 414 POTENTIAL.
FT TRANSMEM 447 467 POTENTIAL.
SQ SEQUENCE 480 AA; 53323 MW; 45799CBD3695B78E CRC64;

Query Match 5.0%; Score 125.5; DB 1; Length 480;
Best Local Similarity 20.1%; Pred. No. 0.078;
Matches 90; Conservative 75; Mismatches 214; Indels 69; Gaps 16;

QY 44 ASAHITAVIGSGVLSLAWIAQAGHIAQSIILLIPSTFTSTMLADCYRAPDPVTGK 103
DB 72 AFMNLANSILGAGITQFFAKNAGILGGLLSYVALGPTVDWTLRLIV---INLTLAGK 127

QY 104 RNYTMDVRSYLGGRKVLQCL---GVAQYGNLIGVTGYTITASISLVAVGKNCNCFHK 159
DB 128 R--TYQGTVEHVMSKKKLLILFTNGLFARGGCI---GYCIIIGDTIPHLRALFQND 181

QY 160 GHTADCTISNYPYNAVGIITQVILSQIPNFKLSPLISMAVMSFTYATIGIGLAITVA 219
DB 182 GNVHFWLRNRIIVVMTFTISPLSMKRNIEALSASFLAVISM-----IIIVLTAVI 234

QY 220 GKVGKTSMTGTAVGV-DVTAAGKIVRSFOAGVDAFAVAYATVLIETQTLRSAPENK 278
DB 235 RGPMLPYDMKSHLKLDFEFWKATIFRSLSVI-SPALVCHNHTSIFPSMNRSV---A 289

QY 279 AMKASLVGVSTTTFFVILCGCTGYAAGNAPGDELDTGFFBFWLIDFANACIAVHL 338
DB 230 KFTLTHISIIISVICCALMGYSGFVFEKTKGNVLSNFPQGTDT--AINIARLCFGFM 347

QY 339 IGAYQVFAQIPQFVEK-----KCN--RNYPNKFTITSEYSVNVVPLGKFNISLPR 388
DB 348 LTFP-----PMEIFVLDVVGNLLHECNLIKVDHTQLSGKHV----- 387

QY 389 VWRAYVYVITTVAMIFPFFNAILGILGASFWPLTVYPPVEMHIAQTKIKKYSARWAL 448
DB 388 VITSLVPIITMGISLITTCNLGALFELIGATTASTWAYILFPYTNLLTSKKK---SW-KE 443

QY 449 KTMCVVCLIVSLAAAGSIAGLISVVK 476
DB 444 RLFPYLCI-----CFGFMIISSQT 465

RESULT 11
FUII_YEAST STANDARD; PRT; 639 AA.
AC P38196;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Uridine permease.
GN FUII OR YBL042C OR YBL0406.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95116707; PubMed=7871888;
```

```
RA de Wergifosse P., Jacques B., Jonniaux J.-L., Purnelle B., Skala J.,
RA Goffeau A.;
RT "The sequence of a 22.4 kb DNA fragment from the left arm of yeast
RT chromosome II reveals homologues to bacterial proline synthetase and
RT murine alpha-adaptin, as well as a new permease and a DNA-binding
RT protein";
RL Yeast 10:1489-1496(1994).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=98146544; PubMed=9485596;
RA Wagner R., de Montigny J., de Wergifosse P., Souciet J.-L., Potier S.;
RT "The ORF YBL042 of Saccharomyces cerevisiae encodes a uridine
RT permease.";
RL FEMS Microbiol. Lett. 159:69-75(1998).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=20408929; PubMed=10827169;
RA "Nucleoside transporter proteins of Saccharomyces cerevisiae.
RT Demonstration of a transporter (FuiI) with high uridine selectivity
RT in plasma membranes and a transporter (FUN26) with broad nucleoside
RT selectivity in intracellular membranes.";
RL J. Biol. Chem. 275:25931-25938(2000).
CC -1- FUNCTION: HIGH-AFFINITY TRANSPORT OF URIDINE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ALLANTOIN PERMEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

EMBL; X78214; CAA5059.1; -.
DR EMBL; Z35803; CAA84862.1; -.
DR PIR; S45776; S45776.
DR SGD; S0000138; FUII.
DR InterPro; IPR001248; Cyt_pur_permease.
DR Pfam; PF02133; Transp_cyt_pur; 1.
DR TIGRFAMs; TIGR00800; ncs1; 1.
KW Transmembrane; Transport.
FT TRANSMEM 163 180 POTENTIAL.
FT TRANSMEM 201 225 POTENTIAL.
FT TRANSMEM 260 276 POTENTIAL.
FT TRANSMEM 284 305 POTENTIAL.
FT TRANSMEM 368 392 POTENTIAL.
FT TRANSMEM 417 435 POTENTIAL.
FT TRANSMEM 461 477 POTENTIAL.
FT TRANSMEM 484 507 POTENTIAL.
FT TRANSMEM 538 562 POTENTIAL.
FT TRANSMEM 573 590 POTENTIAL.
SQ SEQUENCE 639 AA; 72165 MW; 757325DEF4100399 CRC64;

Query Match 4.8%; Score 121.5; DB 1; Length 639;
Best Local Similarity 18.3%; Pred. No. 0.2;
Matches 101; Conservative 80; Mismatches 175; Indels 195; Gaps 28;

QY 35 EKRTGTLWTASAHITAVIGSGVLSLAWIAA---QLG---WIAGTSLILIFSGFITVF-- 85
DB 123 KRRTWTW---KQYIPFWISSFNVN-TWQISATGLQLGNWMTWICWVGYTFVAPFLI 178

QY 86 TSTMADCYRAPDPVTGKKNY-TYMDV-----VRSYLGGRKVLQGVAY 129
DB 179 LGSKVGNNHISPFISRSVFGIYSIWINRVVNVACWNNSTLAYIGSCQVQLMLKAIF 238

QY 130 GNLIIGTVGVTITPASIISLVAVGKSNCFDKGHTADCTISNYPYMA--VFGLIIVLSQIP 187
DB 239 GTNLNTRIKDTI-----KNPNLTNFEFCWVFWVACLPFLWFP 277

QY 188 NFHKLSFLSLMAAVMSFTYATIGIGLATATVAGGK-----VCKTSMGTAVGVDTAAQKI 243


```

Db 278 PDKRHFALSKATPEA---AFGLIWLCKAKGHLALSLNDNGAISKTVLA----- 329
 QY 244 WRSQQAQADIAFAVAATVILIEIOTDLRSSPAENKAMR-----ASLVGVSTTT 292
 Db 330 WSVTRAL--MSALDNEFTLLNAPDTRGKTKYKSSVQSLALPVCYAIISLIGLUSVS 387
 QY 293 FFYLLGCCIGCY-----AAFGNNAPODPLDTGFG-PEBP----- 324
 Db 388 AAYTLYG-VNVMSPDLINRYLDNYSNGRA-GVFLISPIFAFDGLANGSNGSPAGTD 445
 QY 325 ---WLIDFAN-----ACIAVHL-----ICAYVQFQPIF-----QF 352
 Db 446 LTLALPFTIRRGSYICALISAIICPMDLSSSKFTTALAAVAVFLSALAGVISA DYF 505
 QY 353 VEKK-----CNRANYPDNKFTISEYVNVPLGKNISLFLWRTRAYVVTIVVAMIF 405
 Db 506 IVRGYNVIFHCYTDKRGSYVMNKYGTN-----WR-----AVVA YIF 543
 QY 406 ---PEFNAIIGLIG-----AASFMLTYTFPEVEMHIAQTI-- 438
 Db 544 GIADNFAFGFSGSVSPICAMKVYLYNFVGYLLAALSYCILVTFYDIKGIIPDAKITD 603
 QY 439 KKYSARMIATLK 449
 Db 604 KRWLEWVEVE 614

RESULT 12
 YFYA_SCHPO STANDARD; PRT; 591 AA.
 ID YFYA_SCHPO STANDARD; PRT; 591 AA.
 AC 09JUL18;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative amino-acid permease C9.10.
 GN SPAC9.10.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby I., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones W., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moutie S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkert G., Aert R., Robben J., Grymposier B.,
 RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu A., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rocher M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armitage J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: AL121764; CAB57428.1;
 DR InterPro: IPR002293; AA/rel.primase1.
 DR InterPro: IPR004840; AAc.primase.
 DR InterPro: IPR004841; permease.
 DR Pfam: PF00324; aa.primasees; 1.
 DR PROSITE: PS00218; AMINO ACID PERMEASE 1; FALSE NEG.
 KW Hypothetical protein; Transport; Amino-acid transport;
 KW Transmembrane.
 FT TRANSMEM 98 118 POTENTIAL.
 FT TRANSMEM 342 362 POTENTIAL.
 FT TRANSMEM 397 417 POTENTIAL.
 FT TRANSMEM 450 470 POTENTIAL.
 FT TRANSMEM 545 565 POTENTIAL.
 SQ SEQUENCE 591 AA; 65069 MW; FFCDD707AC7849D4 CRC64;
 Query Match 4.8%; Score 121; DB 1; Length 591;
 Best Local Similarity 22.0%; Pred. No. 0.2; Mismatches 194; Indels 120; Gaps 27;
 Matches 108; Conservative 70; Mismatches 194; Indels 120; Gaps 27;
 QY 55 SGV-LSLMAIAQGLWAGTSTLLIFSTFTSTMLADCYRAPDPVTKKNTYMDVVR 113
 Db 94 SGVRLTSWGISFGGPAAYVSMALVGFCSIYTAACLAELCSAL-PAAGS---TYLWAAE 149
 QY 114 SYLG---GRVQVLQGVAGQNGILGVGTITIASISLVAVKSNCFHDKXTACTISNY 170
 Db 150 S-AQPRGRFVRSFL-VAMWSTTKMTTFVASTITOSTANTIFAEVSTFNPPWTND---SDV 204
 QY 171 PYMAVFGI---OVLISQI-PNFHLSFLSMAVMSFTYATIGIATVA--- 219
 Db 205 KFRVQVIVAEVLIVFILLNGVPPRYKMFKAM-LIMFDYVMNIIVPAVATSKPD 263
 QY 220 GGRVQKTSMTGT-----AVGVDTAAQTI---NR---SFGAAGDIAPAY-A 256
 Db 264 GFRSAKVFETFTTYDQAGYKEVDANGNPISLSKIVPKGMQWCLSTFATAGYIVGYDA 323
 QY 259 YATVLIIDOTLSSPAENKAMKRSALVGSITTF-----YILCCIGYAAF--- 306
 Db 324 SGHIAETKDA-----SIKARGIFYSTVTSFIVAFSLALVLFRC-CPDDDTFTAI 373
 QY 307 ---GNNAPODPLDTGFG-----EPFLIDPANACINHLIGAQVRA-----QP 346
 Db 374 LYNNSSQPPVNFVSYLGRGVVNNVITLIEITLNGVVS-LACSLVFAVSRDGLP 432
 QY 349 IFQVEKKCNRNPNKFTISEYVNVPLGKNISLFLWRTRAYVVTIVVAMIF-- 406
 Db 433 FSNMISQVSTKQCKAI-----TVIYISALLICTIPSA 468
 QY 407 -FPAIIGLIGASFMWLTYPPEVEMHIAQTIKYSARMIATLKMTVCVCLIVSLAAG 465
 Db 469 VAFISLVSAAQAPFAAYAVLAFCRFLTRDKPR--GRW-SLGWLSKPCLVITLVY--- 522
 QY 466 SIAGLISVKT Y 477
 Db 523 NLFALVNVSPY 534

RESULT 13
 Y099_MYCTU STANDARD; PRT; 440 AA.
 ID Y099_MYCTU STANDARD; PRT; 440 AA.
 AC Q10858;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical transport protein Rv1999c.

GN RV1999C OR MT2055 OR MTCY39.19.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634220;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaita F.,
RA Badcock K., Baaham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROBABLE AMINO-ACID OR METABOLITE TRANSPORT PROTEIN.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -!- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z74025; CAA98388.1; -;
DR EMBL: AE007057; AAK46332.1; -;
DR TIGR: MT2055; -;
DR TubercuList; Rv1999C; -;
DR InterPro; IPR002295; AA/rel_permease1.
DR InterPro; IPR002422; AA/rel_permease2.
DR InterPro; IPR004840; AAC_permease.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR PROSITE; PS00218; AMINO ACID PERMEASE 1; FALSE NEG.
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 24 44 POTENTIAL.
FT TRANSMEM 47 67 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 155 175 POTENTIAL.
FT TRANSMEM 183 203 POTENTIAL.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 276 295 POTENTIAL.
FT TRANSMEM 323 343 POTENTIAL.
FT TRANSMEM 346 366 POTENTIAL.
FT TRANSMEM 379 399 POTENTIAL.
FT TRANSMEM 400 420 POTENTIAL.
SQ SEQUENCE 440 AA; 45267 MW; 9592EE0A842D18CD CRC64;

Query Match 4.8%; Score 120; DB 1; Length 440;
Best Local Similarity 21.9%; Pred. No. 0.18;
Matches 95; Conservative 52; Mismatches 176; Indels 110; Gaps 20;

OY 24 DPTKNVDEGKEKGTGTLWTASAHITAVIGSGVLSLAWAIAQLGWIAGTSILL----- 77

Db 6 DPRDIPDE--LRRLRLG-LDAVVIGLSMIGAGIFA---ALAPAAAYAGSGLLLGLAVAA 59
OY 78 IFSTFTFTSTMLADCYRAPDPVTGKRNNTYMDVVRSLGKRVQLCGVAQYGNLIGTV 137
Db 60 VVAYCNAISSARLAARY---PASG-GTYVY-----GRMLGDFNGYLA 98
OY 138 GYTTASISLVAVGKNCNCFHDKGHTADC-----TISNYPY-----MAYFGIITQVLSQI 186
Db 99 GWGFWV-----GKTASCAAMALTVGFVWPAQAHAVAVAVVALLTAVNY 142
OY 187 PNFHKLSEFL--SIMAAMSFTVATIGIGLATATVAGGKVGKTSMTGTAVGVVDVTAQKIW 244
Db 143 AGIQKSAWLTRISAVVL-----VVLTAVVAAYGSGAADPARLDIGVDA-----HW 190
OY 245 RSFOAVGDIADAFYA---YATVLIILIOQLASSPAENKAMKASLVGVSITFTFYILCGC 300
Db 191 GMLQAAGLLEFFAFAGYARIATLGEVRDPARTIP-----RAIPLALGITLAVYALVAV 243
OY 301 IGYAAFGNN-----AP-GDFTLDFGFFEPFNLIDFANACIAVHLIGAYQVFAQPIQFV 353
Db 244 AVIAVLGPORLARAAPLSEAMRVAGV---NWLIPVVOIGAAVAALGSLALLILGVSRRT 300
OY 354 EKKCNRYPNPKKITSEYSVNYPFLGKNISLPLRWLTAVYVITTVVAMIFFPFFNAILG 413
Db 301 LAMARDHLPRWLAUAHVHPRKVPF-----RAELVVGAVVAALATAD-IRG 345
OY 414 LIGAASFWEPLTVY 426
Db 346 AIGFSFGLVYVY 358

RESULT 14
YWCA YWCA
ID YWCA BACSU STANDARD; PRT; 516 AA.
AC P39599;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative symporter ywca.
GN YWCA OR IPA-31R.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95020537; PubMed=7934828;
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees.";
RL Mol. Microbiol. 10:371-384 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Bories R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Meliade R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portocarrillo S., Prescott A.M.,
RA Prescan E., Fujie P., Fumelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield P.,
RA Seichuich J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Secorin A., Taccotti A., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi A., Takakoshi A., Tanaka T., Terpietra P., Tomonari A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassarotti A.,
RA Viari A., Wambit R., Wedler E., Wedler H., Weitenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
subtilis."

RL Nature 390:249-256 (1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE SODIUM: Solute symporter family (SSF).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X7124; CABIS187.1; -
CC EMBL: Z9123; CABIS1850.1; -
CC Subtilist; BG10577; YWCA.
DR InterPro: IPR001734; Na/solut_symport.
DR Pfam: PF00474; SSF; 1.
DR TRIGRAMS: TRIGR00813; ssa; 1.
DR PROSITE: PS00456; NA_SOLUT_SYM_1; 1.
DR PROSITE: PS00457; NA_SOLUT_SYM_2; 1.
DR PROSITE: PS00283; NA_SOLUT_SYM_3; 1.
KW Hypothetical protein; transport; Transmembrane; Sodium transport;
KW Symport; Complete proteome.
FT TRANSMEM 7 27 POTENTIAL.
FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 74 94 POTENTIAL.
FT TRANSMEM 132 152 POTENTIAL.
FT TRANSMEM 156 176 POTENTIAL.
FT TRANSMEM 180 200 POTENTIAL.
FT TRANSMEM 236 256 POTENTIAL.
FT TRANSMEM 276 296 POTENTIAL.
FT TRANSMEM 323 343 POTENTIAL.
FT TRANSMEM 377 397 POTENTIAL.
FT TRANSMEM 398 418 POTENTIAL.
FT TRANSMEM 427 447 POTENTIAL.
FT TRANSMEM 465 485 POTENTIAL.
SQ SEQUENCE 516 AA; 54958 MW; E33628B26A3952D CRC64;

Query Match 4.7%; Score 118.5; DB 1; Length 516;
Best Local Similarity 19.3%; Pred. No. 0.26;
Matches 93; Conservative 73; Mismatches 216; Indels 100; Gaps 19;

OY 262 ----VLEIETDLTRSSPAENKRAMKASLVGVSTTFYILGCGIGY--AAF----- 306
DB 250 GLPHILIRFTYVDATATATSVASATIGV-----FTIMVFIFGFAAFVGFDAITAA 304
OY 307 ---GNNAF-----GDFLDFGFEPEPMLIDPAN--ACIAVLIAQVQFAPQIFQV 353
DB 305 DQAGNMAAPLAKALOGDPL--FAFVS--IAFETILAVTGLVLSAAGAFHDIYSQI 359
OY 354 EKCNNRRYPPNKRFTTSEYGVNPEFGKNISLFLWETAVVITTVAMIF-----PF 407
DB 360 IRGCEATEK-----QMKARMAVSVAVSLSLILAFQASLVNAP 399
OY 408 FNAILGLIGASFPPLTVYPVEKMIADTKIKYSARNIALKTCVCLIVSLIAAGST 467
DB 400 LVLAFAVAASANLPLIVP-----TVFWKRFNASCALMGSLTGLISALVIVSMSPVWDP 454
OY 468 AG 469
DB 455 AG 456

RESULT 15
YSPK CAEEL STANDARD: PRT; 615 AA.
ID YSPK CAEEL
AC 019425;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein F13H10.3 in chromosome IV.
GN F13H10.3.
OS *Caenorhabditis elegans*.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; *Caenorhabditis*.
OX NCBI_TaxID=6239;
KP SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Cottage A.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.; (OCT-2001) to the EMBL/GenBank/DBJ databases.
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: SOME, TO YEAST YEL064C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z68748; CAA32953.2; -
DR WormPep; F13H10.3; CE27359.
DR InterPro: IPR002422; AA/re1_pmease2.
DR Pfam; Pf01490; Aa_trans; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 195 215 POTENTIAL.
FT TRANSMEM 248 268 POTENTIAL.
FT TRANSMEM 342 362 POTENTIAL.
FT TRANSMEM 368 388 POTENTIAL.
FT TRANSMEM 416 436 POTENTIAL.
FT TRANSMEM 452 472 POTENTIAL.
FT TRANSMEM 496 516 POTENTIAL.
FT TRANSMEM 536 556 POTENTIAL.
FT TRANSMEM 562 582 POTENTIAL.
FT TRANSMEM 595 615 POTENTIAL.
SQ SEQUENCE 615 AA; 68692 MW; 5092F2A8F4A3A420 CRC64;

Query Match		4.6%	Score 115.5;	DB 1;	Length 615;
Best Local Similarity		18.5%	Pred. No. 0.52;		
Matches		90;	Conservative 81;	Mismatches 164;	Indels 151; Gaps 21;
QY	48	IITAVIGSGVLSLAWIAQAGT	AGT	SILLIFSFITYFTSTMLADCYR	-----APDPVTG 102
Db	172	IWNWMTSLSLAMPWALQQAG	LVLCIIILMSMAAICFYAYIVIESPR	KLQDLSVDPLLA 231	
QY	103	KRNYTMDVVRSYLGGKRVOLC	-----GVAQYGNLI	-----GVTGVYTTIAS 144	
Db	232	E-----FSDVCKS-LFGRIGEY	CAVFSVCVLIGGVIYVWVLM	SNFLYVTGAVVYESMQPN 286	
QY	145	ISLVAVGKSNCF	-----	HDKGHTADCTISNVPYMAVF 176	
Db	287	STTIPVMENKFTCDVVC	PGTSQWTIPQWEKOLYDAVSEMEGG	TGDDSHSFDKFWTLR 346	
QY	177	GIQVILS-----QIPNPHKLSF	-----LSIMAAMVSFTYATIGIGLA	IATVAGKV 223	
Db	347	GTVPYILAFALPFLMNEKSP	FTFTKFNVLGTISVMYLLM-FVF	-----	388
QY	224	GKTSMTGTAVGVDTAAQKIW	-----RSFOAV-GDIAFAYAVATV	LIEIQDTLRSSP-A 275	
Db	389	--SKLECGVNDFSNPKSIHY	VOLAMHPALSGTILSYFIHNAV	TI-----LRNQHP 443	
QY	276	ENKAMKRASLVGVSTTTFFYL	CGCIGYAAF--GNNAPGDFLTD	FGFPFWLIDFANA 332	
Db	444	ENNA--RDLSIGYCLVAF	CVYVIGFTFFAAFPVQRSCISD	NFLNFG-----AGDVLSS 495	
QY	333	CIAVHLIGAYOVFAQPI	FPQVPEKKCNRYPDNKFITSEY	SVNVVPLGKFNISLRLVWRT 392	
Db	496	TARLELL-FQMITVLP	LLMLVR-----	SQLFYAIFCQTHPG 531	
QY	393	AYVVTITVYVMI-----	FPFFNAILGLIGAASFWEPLTVY	FPVEMHIAQ-----T 436	
Db	532	AIRVILNVLIIAVAVGFAT	FPNPGSILRYVGSISGLVYV	FALPAMVYIKQSEAAAGTLT 591	
QY	437	KIKKYS 442			
Db	592	PMKKYA 597			

Search completed: November 22, 2002, 13:30:18
Job time : 17 secs